

UNIVERSIDAD NACIONAL DE LA PLATA



ADAPTABILIDAD CONFORMACIONAL DE LAS HÉLICES CENTRALES DE APOLIPOPROTEÍNA A1 HUMANA

SIMÓN MONGI

TRABAJO FINAL

PARA OBTENER EL TÍTULO DE

LIC. EN BIOTECNOLOGÍA Y BIOLOGÍA MOLECULAR

DIRECCIÓN:

DR. PROF. MCCARTHY, ANDRÉS NORMAN

TUTORÍA:

LIC. SOLEDAD LYNN

LA PLATA, BUENOS AIRES, ARGENTINA

2016

*En el trajín de vivir vamos aprendiendo a ser.
Y no hay mucho más que decir, por suerte hay mucho que hacer.
Me empujaste sin tocarme, en busca de mi destino,
y hoy te envío una postal de donde he venido a parar.*

Científicos del Palo - Lo que el viento nos dejó

Para mi madre, **Beatriz**,
que gracias a su paciencia, esfuerzo y dedicación pude lograr esta meta.

Agradecimientos

A mis amigos de la vida; Santy, Agus, Dito, Lucho, Nacho, Mauro y Barri, que a pesar de todo estuvieron conmigo durante toda la carrera, y aceptaron cada vez que no pude estar para ellos.

A Luciano, con el que supimos manter una amistad invaluable, a la distancia y no tan lejos.

A mis amigos de la facultad;

A Gaston por ser un amigo incondicional, y por dar valiosos consejos.

Juli, Cande y Diana que comparitmos desde el ingreso juntos, gran parte de la carrera, y que sin ellas el peso de la facultad hubiese sido una carga mucho mas pesada de lo que fué.

Diana, Corel, Eliana, Martina y Guillermo con las que compartimos muchísimas horas de cursada y estudio.

A mi director, Andres, y a mi tutora, Sole, por su tiempo y dedicación para la enseñanza de temas que excedieron los academicos.

A mis compañeros del CREG; Jor y Rulo, que por ellos la realización de este trabajo fue posible y mucho mas divertido.

A los compañeros de Suma, de todos los claustros, que me enseñaron a dimensionar a la facultad y a nuestro conocimiento en un contexto social, a como transformarla, a ser irreverente.

A los compañeros del PEAS, del Banco de Germoplasma y de Extension.

A la facultad de Ciencias Exactas, a la institución y a todas las personas que la transitamos.

Índice general

1. Introducción	1
1.1. Aterosclerosis y el colesterol	1
1.2. Colesterol: síntesis y metabolismo en humanos	2
1.3. Transporte reverso de colesterol (RCT)	4
1.4. Apolipoproteína A1: Estado del Arte	5
2. Objetivos	11
2.1. Objetivos generales	11
2.2. Objetivos específicos	11
3. Materiales y métodos	13
3.1. La dinámica molecular	13
3.1.1. ¿Qué es?	13
3.1.2. Aproximaciones	14
3.2. Principios de la simulación	16
3.3. Campos de fuerza	17
3.4. Mecánica molecular	17
3.4.1. Visual Molecular Dynamics (VMD)	17
3.4.2. CELLmicrocosmos 2.2 MembraneEditor	18
3.5. Armado del modelo de partícula LL5/2	18
3.6. Reconstrucción del modelo LL5/5	21
3.7. Armado de la bicapa lipídica	22
3.8. Compatibilización del dímero proteico con la bicapa lipídica . . .	23
3.9. Preparación del sistema para las simulaciones	25
3.10. Estabilización	26
3.11. Adecuación de campo de fuerza a la bicapa elegida	27

4. Resultados y discusión	29
4.1. Detalles de las simulaciones	30
4.2. Análisis de las simulaciones	30
4.2.1. RMSD	30
4.2.2. Análisis de superficie accesible al solvente (SASA)	34
4.2.3. Análisis de estructura secundaria	36
4.2.4. Cálculos de deuterium order parameters	39
4.2.5. Distancias entre residuos	42
5. Conclusiones	57
Bibliografía	59

Capítulo 1

Introducción

1.1. Aterosclerosis y el colesterol

Las enfermedades cardiovasculares continúan siendo una de las principales causas de muerte en el mundo, correspondiendo al 31% de las muertes totales a nivel mundial en el 2008.

De todas estas, la aterosclerosis es responsable de cerca del 80% de estas muertes. Esta es una enfermedad compleja, que se desarrolla en las paredes de los vasos sanguíneos a través de los años, donde se generan depósitos de material graso y colesterol dentro del lumen de arterias medianas o grandes, lo que provoca que la superficie interna de los mismos se vuelva irregular y se reduzca el área por donde circula la sangre. Eventualmente, estos depósitos, llamados placas, pueden romperse, generando la formación de un coagulo, que dependiendo de la arteria donde se desarrolle puede resultar en infartos o en accidentes cerebrovasculares (ACV) [Mendis et al., 2011].

Se conoce desde hace un tiempo, que la proteína principal de las lipoproteínas de baja densidad (HDL), la Apo A1, es un factor

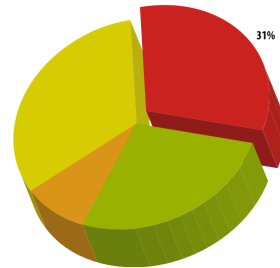


Figura 1.1:
Causas de muerte en el mundo. Rojo corresponde a enfermedades cardiovasculares.

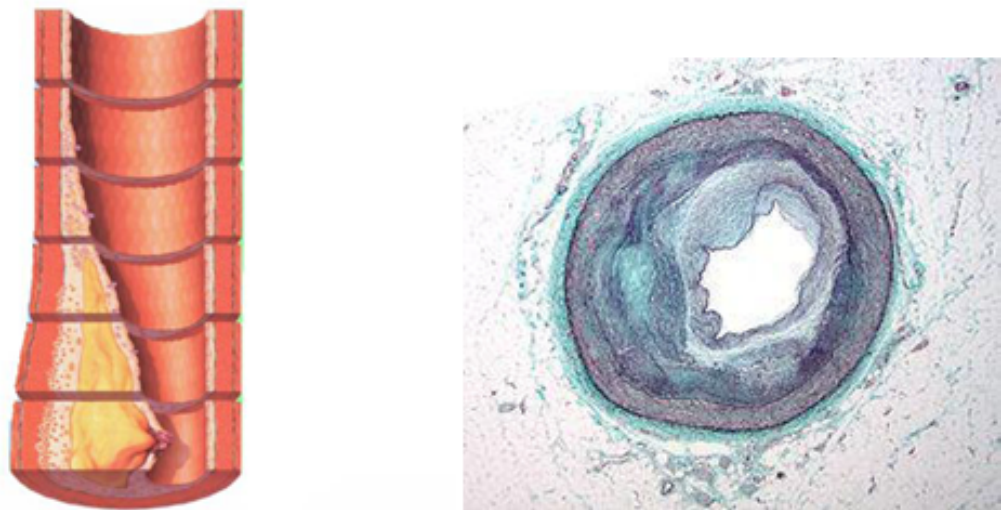


Figura 1.2:

Aterosclerosis. A la izquierda se ve un esquema de acumulación de placa a través del tiempo. A la derecha un preparado de corte transversal de arteria con placa.

de riesgo negativo para enfermedades cardiovasculares [Wilson et al., 1988, Vergeer et al., 2010], así como también son los niveles altos de HDL plasmático [Miller, 1987, Vergeer et al., 2010]. Además, se ha demostrado que la sobreexpresión de APO A1 en animales transgénicos llevo a la inhibición de aterosclerosis, reforzando la idea de su rol protector [Segrest et al., 2000].

1.2. Colesterol: síntesis y metabolismo en humanos

El colesterol es un lípido esteroide, con una cabeza polar constituida por un grupo hidroxilo, y una cola apolar formada por 4 carboxilos condensados y sustituyentes alifáticos. Posee muy baja solubilidad en agua, siendo un componente importante de las membranas plasmáticas de las células animales, cuya función es regular su fluidez y también es precursor de la vitamina D, hormonas y sales biliares. El colesterol puede ser ingerido en la dieta o puede ser sintetizado de Novo en las células a partir de acetil-CoA. Su síntesis se encuentra

fuertemente regulada, dependiendo de los niveles de colesterol intracelular, y de los niveles ingeridos, para evitar su acumulación. La eliminación del colesterol ocurre a través de la bilis como colesterol libre o como sales biliares, estas posteriormente se convierten en ácidos biliares, para ser degradados por las bacterias intestinales y finalmente expulsados en las heces. Debido a su baja solubilidad, el transporte en sangre no se da en forma libre sino a través de lipoproteínas y fosfolípidos. Cinco partículas diferentes fueron nombradas según su separación en gradientes de densidad y también según destino celular. Los quilomicrones son las partículas encargadas de transportar los lípidos desde el intestino delgado al organismo a través del sistema linfático y circulatorio. Luego de este recorrido por el organismo, el colesterol de los quilomicrones que no ha sido captado en los tejidos es absorbido por el hígado, donde ejerce su acción regulatoria sobre la síntesis endógena.

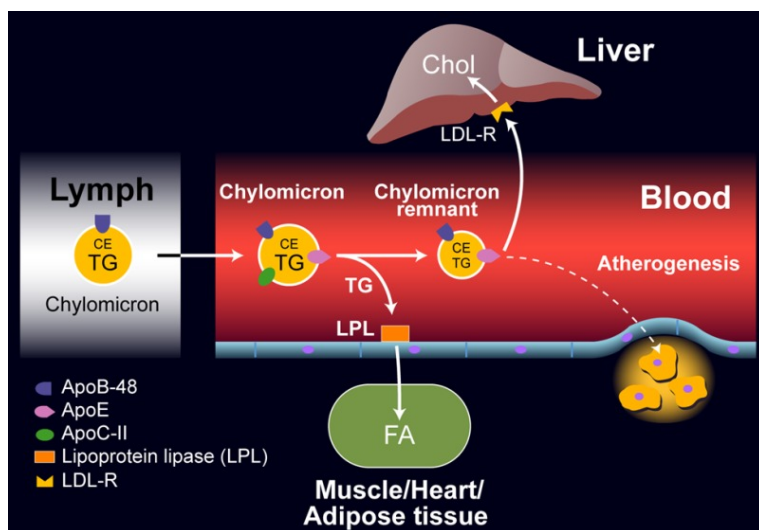


Figura 1.3:

Esquema de ingesta de lípidos y lipoproteínas involucradas. Grafico extraído de [Feingold and Grunfeld, 2015]

El colesterol que se encuentra en exceso en el hígado, se transporta en el suero dentro de las LDLs (Low Density Lipoprotein). El hígado sintetiza VLDLs (Very Low Density Lipoprotein) y éstas se convierten a IDL (lipoproteínas de

densidad intermedia) y sus remanentes a LDLs.

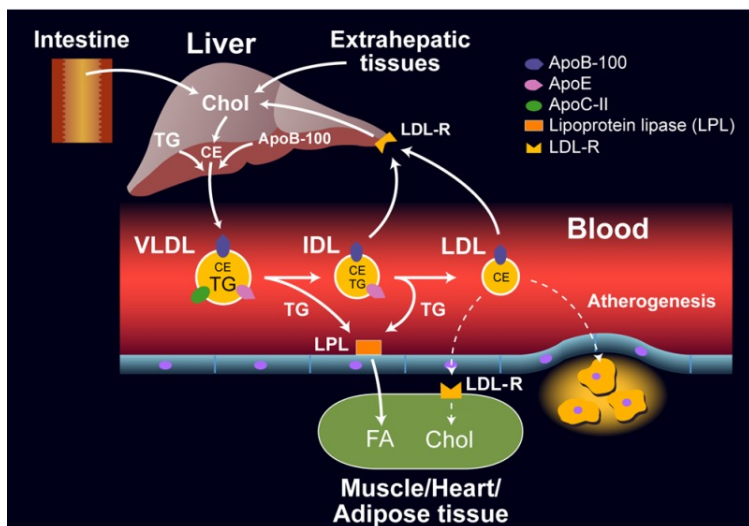


Figura 1.4:

Esquema de circulación de lípidos endógenos y lipoproteínas involucradas.
Grafico extraído de [Feingold and Grunfeld, 2015]

Por último, las HDL (High Density Lipoproteins) son las encargadas del llamado transporte reverso de colesterol. Mediante este proceso, el colesterol es captado de los tejidos y de sus depósitos en las arterias, y es transportado al hígado para su eliminación.

En resumen el colesterol que se produce, y el procedente de los remanentes de los quilomicrones es recogido por el hígado. Desde allí se transporta en las lipoproteínas VLDL, IDL y LDL principalmente a los tejidos para su uso, mientras que el colesterol que se transporta desde los tejidos al hígado para ser eliminado es transportado en HDL [Kwiterovich, 2000, Feingold and Grunfeld, 2015].

1.3. Transporte reverso de colesterol (RCT)

Las HDL son un grupo de partículas heterogéneas, siendo la Apolipoproteína A-I (ApoA1) el principal componente proteico, que constituye cerca del 70% del total. El proceso llamado transporte reverso de colesterol involucra a

estas partículas, y se ha dilucidado a nivel molecular, diferenciándose en 3 etapas. En primer lugar, los fosfolípidos son translocados por el transportador ABCA1 desde la hemicapa citosólica hacia la hemicapa periplásmica, protruyendo la membrana de una manera similar a la formación de vesículas. La proteína ApoA1 libre es reclutada para generar HDL nacientes discoidales (nHDL) ricas en fosfolípidos. Estas nHDL contienen una bicapa lipídica conteniendo colesterol y rodeado por las alfa hélices de la ApoA1. Las nHDL catalizan el eflujo de colesterol adicional desde la membrana plasmática a la HDL naciente. En segundo lugar, en la circulación general, la activación de la lecitina:colesterol aciltransferasa (LCAT) convierte las moléculas de colesterol libre (FC) en ésteres de colesterol (CE). Este colesterol esterificado, se mueve al centro de la bicapa, llevando a la maduración de las nHDL y a la formación de partículas esféricas llamadas HDL3. Estas pequeñas partículas esféricas ricas en colesterol esterificado son otra vez remodeladas por la acción de LCAT y otros factores plasmáticos, como la proteína de transferencia de ésteres de colesterol (CETP), formando la fusión de HDL3 para generar HDL2. Estas partículas de mayor tamaño liberan a la ApoA1, para ser reutilizada o degradada. Las HDL3 y HDL2 son las partículas más abundantes encontradas en circulación, y consisten en un monocapa de lípidos junto a proteínas anfipáticas en la superficie, y ésteres de colesterol junto a algunos triglicéridos en su interior hidrofóbico. Por último, la tercera etapa consiste en la interacción de estas partículas con receptores hepáticos, como el receptor Scavenger B, tipo 1 (SR-B1), que lleva a la captación de estos ésteres de colesterol por el hígado y su posterior excreción a través de la bilis. Esta captación, promueve la disociación de ApoA1, que vuelve a la circulación como ApoA1 libre, donde puede ser reciclada, o degradada (Figura 1.3).

Esta habilidad de las HDL de movilizar el colesterol para su eliminación a través del transporte reverso de colesterol explica, en parte, su rol de protección frente a la aterosclerosis [Jones et al., 2009, Guha et al., 2008, Gursky, 2015].

1.4. Apolipoproteína A1: Estado del Arte

La Apolipoproteína A1 es una proteína humana de 243 aminoácidos y es codificada por los exones 3 y 4 del gen Apolipoproteína en el cromosoma 11. El exón 3 codifica para los aminoácidos 1 al 43, que conforma el llamado dominio globular. Este contiene al extremo N-terminal junto a 10 residuos, y seguido de

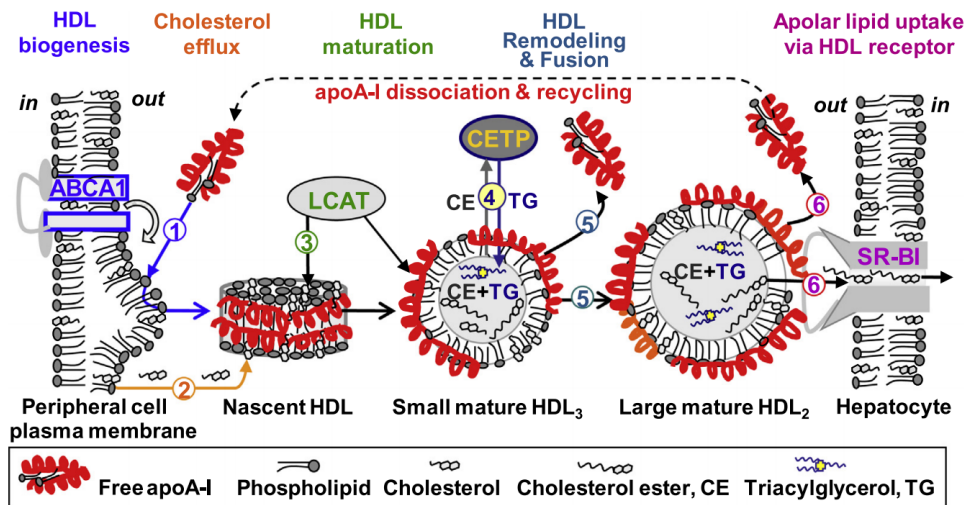


Figura 1.5:

Esquema de transporte reverso de colesterol. (1) Apo-A1 interactúa con la membrana plasmática y el transportador ABCA1 para generar la partícula HDL discoidal (2) Estas partículas toman colesterol adicional de la membrana (3) donde la LCAT lo esterifica, generando un núcleo hidrofóbico. (4) Las partículas se siguen remodelando por acción de LCAT y otros factores, produciendo un desbalance en la superficie y el núcleo, y (5) liberando Apo-A1 que se recicla o se degrada (6) a medida que madura la partícula [Gursky, 2015].

3 repeticiones en tándem de 11 residuos. El exón 4 codifica para el llamado dominio asociado a lípidos (lipid-associating domain), el cual contiene 8 repeticiones en tándem de 22 residuos cada uno y 2 repeticiones de 11 residuos cada uno, con una estructura secundaria predominante de α -hélices antipáticas [Segrest et al., 2000, Mei and Atkinson, 2011].

La información estructural a escala atómica es limitada, debido a que la ApoAI tiene conformaciones flexibles, y resulta difícil preparar soluciones homogéneas. Esta se puede encontrar en múltiples composiciones y en partículas de tamaños distintos, así como también distinta cantidad y tipos de lípidos. Es-

ta dificultad ha llevado a proponer una multitud de modelos de nHDL a través de los años [Gogonea, 2016]. Nos concentraremos en modelos discoidales de nHDL, los cuales muchos autores consideran suficientes para explicar su gran diversidad estructural y funcional [Phillips, 2013].

Modelos propuestos:

- Picket Fence (valla de estacas)

Modelo construido a partir de una bicapa lipídica de 144 POPC, 12 coles-teroles libres (FC) y un dímero de ApoAI arreglados en una conformación de cabeza con cabeza, orientado con sus α -hélices de manera similar a las transmembrana. Posteriormente este modelo fue refinado por simulaciones de MD. El modelo explica los resultados obtenidos en estudios de espectroscopia infraroja (IR) y SANS (small angle neutrón scattering). Figura 1.6 A.

- Double belt (cinturón Doble)

Este modelo cobró fuerza a partir de la obtención de la estructura cristalo-gráfica de la ApoAI sin lípidos, truncada en su N terminal (43–ApoAI). Como se ve en la figura 1.6 B, un dímero de ApoAI se encuentra uno sobre el otro de manera antiparalela, y sus dominios hidrofóbicos están interaccionando con una bicapa de POPC, rodeándola. Este modelo responde además a estudios de EM que mostraba a las nHDL como partículas que se “apilan como monedas”, y, a posteriori, a una variedad de estudios biofísicos y de simulaciones de dinámica molecular. En la imagen se observa el modelo propuesto por J. Segrest [Segrest et al., 1999], donde las Delta-43–ApoAI se encuentran antiparalelas, una encima de la otra, coincidiendo su hélice 5 (modelo LL5/5).

- Hairpin (horquilla para pelo)

Este es un modelo alternativo propuesto por Silva [Silva et al., 2005], donde las dos cadenas de ApoAI no están asociadas de manera antiparalela entre sí sino que forman independientemente una semi circunferencia y únicamente están asociadas de forma inter-catenaria por solo un pequeño grupo de aminoácidos.

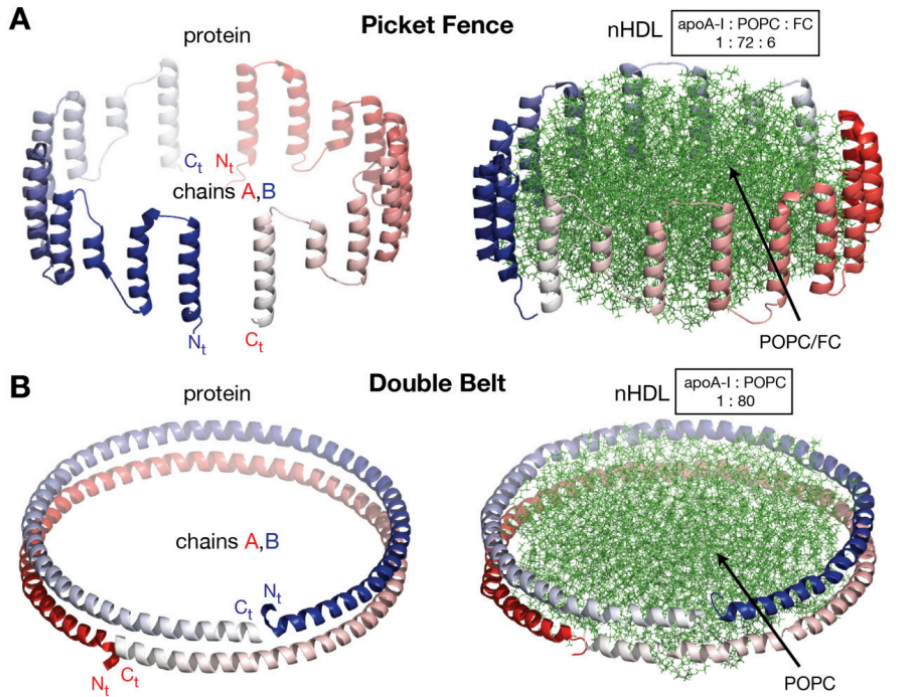


Figura 1.6:

Modelos discoidales:

A Izq: Dimer de Apo-A1 en el modelo de valla de estacas.

A Der: modelo de valla de estacas en una bicapa de 144 POPC y 12 FC.

B Izq: Dimer de Apo-A1 en el modelo computacional de cinturón doble.

B Izq: modelo de cinturón doble con una bicapa de 160 POPC.

Figura extraída de [Gogonea, 2016]

- Looped Belt (cinturón con bucle)

En este modelo, Lagerstedt [Lagerstedt et al., 2012], propone una variación del modelo del cinturón doble, a partir de que varios estudios revelaran que ciertas regiones de la ApoAI son altamente dinámicas. En este, la región de Apo-AI comprendida por parte de la hélice 5 y el inicio de la hélice 6, se encuentra en una conformación random coil (serpentina aleatoria), siendo este un dominio clave para la funcionalidad de la

ApoA1.

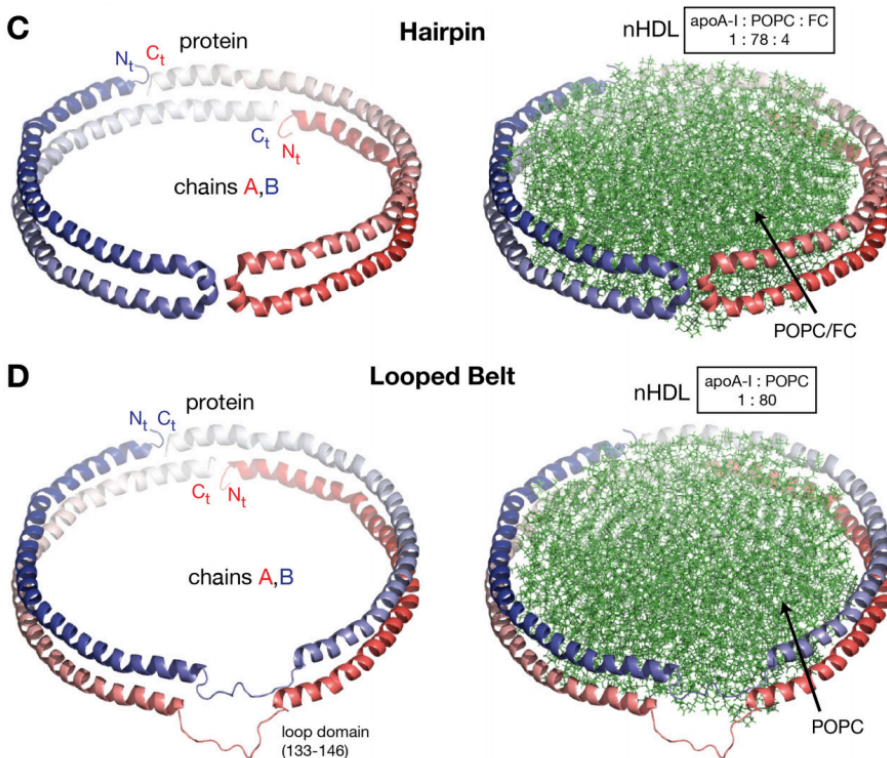


Figura 1.7:

*Modelos discoidales:***C** Izq: Dimero de Apo-A1 en el modelo de horquilla para pelo.**C** Der: modelo hairpin embebido en una bicapa de 156 POPC y 8 FC.**D** Izq: Dimero de Apo-A1 en el modelo de cinturón con bucle.**D** Izq: modelo de cinturón con bucle con una bicapa de 160 POPC.

Figura extraída de [Gogonea, 2016]

- Solar Flares (ráfagas solares)

A partir del modelo Looped Belt, pero incorporando información a partir de otros estudios, Wu [Wu et al., 2007], propone este modelo, el cual consiste en una ApoAI de secuencia completa. Similar al modelo Double belt, como se observa en la 1.8 E el dímero de ApoAI se encuentra orientado de manera antiparalela y forman un anillo, alrededor de una bicapa lipídica compuesta de 200 POPC y 20 colesteroles libres. Además, se encuentra añadido el dominio N terminal, el cual se encuentra en una conformación globular.

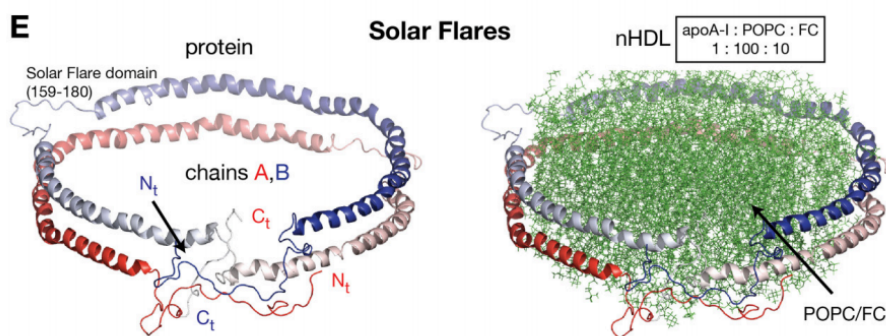


Figura 1.8:

Modelos discoidales:

E Izq: Modelo computacional de Apo-A1 en el modelo de ráfagas solares.

E Der: modelo de ráfagas solares en una bicapa de 200 POPC y 20 FC.

Figura extraída de [Gogonea, 2016]

A pesar del hecho de que estos modelos incorporan enormes datos de estudios biofísicos, ninguno está basado en observaciones directas ni de la ApoAI ni de los lípidos. Por lo tanto, la conformación de la ApoAI se encuentra todavía en debate.

Por fuera de estos modelos, se han presentado nuevos modelos, derivados de estudios realizados con SANS con variación de contraste. Modelos como el Double Superhelix, Turtle o Horseshoe han encontrado gran resistencia en su aceptación [Gogonea, 2016].

Capítulo 2

Objetivos

2.1. Objetivos generales

- Avanzar en el conocimiento de la organización estructural de la Apo-AI en HDL discoidales.

2.2. Objetivos específicos

- Preparar y estabilizar el modelo de cinturón doble de HDL discoidal con Apo-AI en los registros LL5/5 y LL5/2, mediante el uso de métodos computacionales de mecánica molecular.
- Adecuar el campo de fuerza ff43a1-S3 de lípidos para su uso en el programa GROMACS de dinámica molecular.
- Estudiar mediante técnicas de dinámica molecular la estabilidad conformacional de la Apo-AI en los registros antes mencionados.
- Comparar los modelos obtenidos mediante dinámica molecular con datos experimentales de bibliografía.

Capítulo 3

Materiales y métodos

3.1. La dinámica molecular

3.1.1. ¿Qué es?

Las simulaciones por Dinámica Molecular consisten en un método de cálculo numérico mediante el cual se resuelven las ecuaciones de la dinámica de Newton para un sistema de N partículas que interaccionan entre sí. Las ecuaciones se resuelven simultáneamente e iterativamente en pasos de tamaño establecido. Es decir, el sistema evoluciona durante algún tiempo, manteniendo la temperatura y la presión dentro de un rango estipulado alrededor de los valores requeridos, y luego del tiempo establecido, se registran las coordenadas atómicas y se vuelven a resolver las ecuaciones, dejando evolucionar al sistema otra fracción de tiempo establecido.

Luego de los cambios iniciales, el sistema normalmente tenderá a alcanzar un estado de equilibrio, en el que minimice su energía. Muchas propiedades macroscópicas pueden extraerse de los datos de salida, cuando se promedia sobre una trayectoria de equilibrio. En otros casos, se busca que las simulaciones se realicen fuera del equilibrio, para obtener otra información.

Si queremos analizar el comportamiento de alguna molécula o partícula con esta técnica, primero debemos desarrollar un modelo de la misma, realizar la simulación y luego contrastar los datos de salida con medidas empíricas. Este modelo puede ser nutrido con estructuras cristalográficas o de modelado computacional, y ser modificado por medio de mecánica estructural a partir

de información bibliográfica. De ser validado, nuestro modelo de escala atómica detallada, podrá predecir propiedades macroscópicas que, a posteriori, planteará nuevas interrogantes y permitirá diagramar posibles estrategias para abordarlas. Aún así, es apropiado considerar las aproximaciones que utilizan estas simulaciones para entender el alcance de las mismas.

3.1.2. Aproximaciones

- Las simulaciones son Clásicas

Como hemos mencionado, la utilización de las ecuaciones de movimiento de Newton implica automáticamente el empleo de la “Mecánica Clásica” para la descripción del movimiento de los átomos. Es decir, en estas simulaciones no se toman en cuenta interacciones de carácter esencialmente mecánico-cuántico. Por ejemplo, en reacciones químicas donde los enlaces se debilitan, se rompen o se crean; o también, para el caso de los átomos de hidrógeno, cuya masa al ser tan pequeña, puede moverse a través de una barrera de potencial por efecto túnel, conformando una transferencia de un protón por medio de un puente de hidrógeno. Aún así, la aproximación clásica resulta adecuada para la descripción de movimiento de la mayoría de los átomos a temperaturas normales. Es decir, en estas simulaciones, los enlaces covalentes y sus ángulos se tratarán con restricciones topológicas, como por ejemplo, en el caso de los hidrógenos, éstos se adsorben al átomo pesado al que están unidos [Lindahl et al.,].

- Los electrones se encuentran en el estado energético fundamental

En estas simulaciones, utilizamos la aproximación de Born-Oppenheimer, la cual supone que los electrones ajustan su dinámica instantáneamente al cambiar las posiciones atómicas. Es decir, que los movimientos electrónicos no son tomados en cuenta, y además, permanecen en un estado cuántico no excitado. Esta aproximación no puede dar respuesta a, por ejemplo, los procesos de transferencia electrónica, por lo que las reacciones químicas en general no se observarán mediante ésta técnica, aunque pueden estudiarse localmente utilizando algunas herramientas [Lindahl et al.,].

- El Campo de Fuerza

El Campo de Fuerza es un conjunto de ecuaciones usadas para generar los potenciales de interacción interatómica y sus parámetros. A partir de ecuaciones y junto a las ubicaciones atómicas, se calculan las fuerzas que gobernarán a ese sistema en ese instante de tiempo. Todos los campos de fuerzas disponibles poseen limitaciones, y estos no son verdaderamente parte del método de Dinámica Molecular. La existencia de un gran número de campos de fuerza permite la utilización de diferentes parametrizaciones de acuerdo al tipo de sistema a estudiar. Son susceptibles de modificaciones según se necesite, o bien se actualicen los conocimientos en la materia. Estos toman en cuenta 3 tipos de funciones potenciales: bonded (enlazantes), non-bonded (no enlazantes), y restricciones. Las interacciones enlazantes, son tratadas como una lista fija de átomos, donde se incluyen interacciones de a 2, 3 o 4 puntos, como enlaces, ángulos y diedros. Las no enlazantes se calculan de a pares, es decir, con un término de repulsión y otro de atracción. Las interacciones no enlazantes que no pueden ser descritas por interacciones de a pares, son representadas por potenciales efectivos de a pares, es decir, contribuciones promedio, las cuales fueron calculadas en determinadas condiciones e integradas a los potenciales. Es necesario entonces tener en cuenta en qué condiciones fueron parametrizadas estas contribuciones para que se verifiquen en nuestro sistema. Las restricciones incluyen datos experimentales o corregir desviaciones [Lindahl et al.,].

- Las interacciones de largo alcance son tratadas con radios de corte

En general en los trabajos de Dinámica Molecular, los cálculos con el campo de fuerza se volverían prohibitivos si se aplicaran a sistemas completos con gran cantidad de partículas. Por lo tanto, se utilizan métodos para el tratamiento de interacciones de largo alcance, tanto para las interacciones de Lennard-Jones como para las interacciones Coulómbicas. Uno de estos métodos es el de radios de corte o cut-off y otro el suavizado o shift. El cut-off considera un radio de corte que trata al sistema como un medio continuo, de permitividad dieléctrica constante. El shift se combina con el anterior y realiza un suavizado en las funciones truncadas por el radio de corte del método cut-off. Estos dos métodos son tratados en diferentes esquemas, utilizando diferentes parámetros

[Lindahl et al.,].

3.2. Principios de la simulación

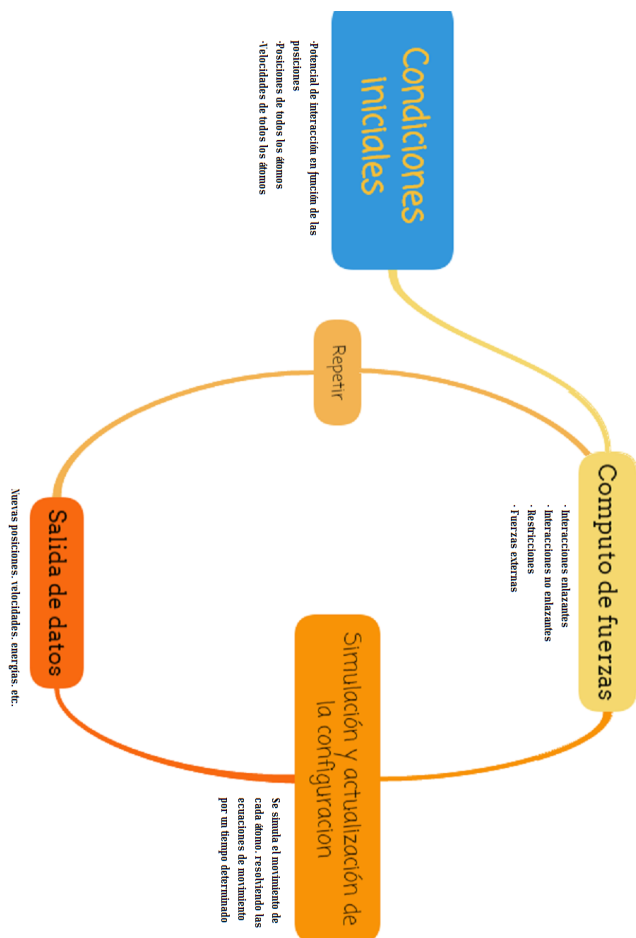


Figura 3.1:
Esquema de simulación. Modificado y traducido del manual del paquete GROMACS [Lindahl et al.,]

3.3. Campos de fuerza

El paquete de programas GROMOS, abreviatura de GRONingen MOlecular Simulation (Simulación Molecular Groninga) contiene el campo de fuerza utilizado para MD del mismo nombre. Este se ha desarrollado en continuo desde 1978 por W. F. van Gunsteren y colaboradores, que es utilizado para el modelado dinámico de biomoléculas a través de numerosas técnicas. En particular, en este trabajo se usó el GROMOS96, en el cual se reescribe completamente los algoritmos de MD, y minimización de energía (EM) a su predecesor, GROMOS87, para incorporar mejoras y nuevas funcionalidades [van Gunsteren et al., 1996, Scott et al., 1999]. El conjunto de parámetros que se utilizaron fue el 43a1, el cual es el indicado para proteínas en solución acuosa [Best et al., 2008].

El campo de fuerza 43a1 de GROMOS96 no contempla a lípidos y otras macromoléculas de interés. En estos casos es donde se debe recurrir a otros campos de fuerza que tengan parametrizadas estas moléculas. El campo de fuerza 43A1-S3 es un campo de fuerza utilizado para la simulación de membranas lipídicas por el paquete Gromacs. El campo de fuerza GROMOS96 43A1, utilizado como base, fue reparametrizado completamente de una manera progresiva, basándose en la reproducción de volúmenes moleculares y la reproducción de calores de vaporización de los lípidos, y por lo tanto, reproduce adecuadamente las moléculas de DMPC y Colesterol en modelos de bicapa [Chiu et al., 2009].

3.4. Mecánica molecular

3.4.1. Visual Molecular Dynamics (VMD)

VMD es una sigla que en inglés significa Dinámica Molecular Visual (Visual Molecular Dynamics). Es un programa de visualización molecular, que se encarga de mostrar, animar y analizar sistemas biomoleculares, y representándolos en 3 dimensiones. Estos sistemas biomoleculares pueden ser proteínas, ácidos nucleídos, bicapas lipídicas y moléculas en general. Además, posee numerosos plugin y una interface de scripts para aumentar su funcionalidad y versatilidad [Eargle et al., 2006, Frishman and Argos, 1995, Humphrey et al., 1996, Sanner et al., 1995, Sharma et al., 2000, Stone et al., 2001, Stone, 1998, Varshney et al., 1994].

3.4.2. CELLmicrocosmos 2.2 MembraneEditor

El programa CELLmicrocosmos 2.2 MembraneEditor (CmME) realiza un enfoque modular de estructuras PDB, ya sea lipídicas o proteínicas, donde se utilizan algoritmos basados en la forma de las moléculas y sus radios de van der Waals para generar modelos de membrana. Luego la membrana puede ser exportada para ser utilizada con los paquetes de simulación [Sommer et al., 2011b].

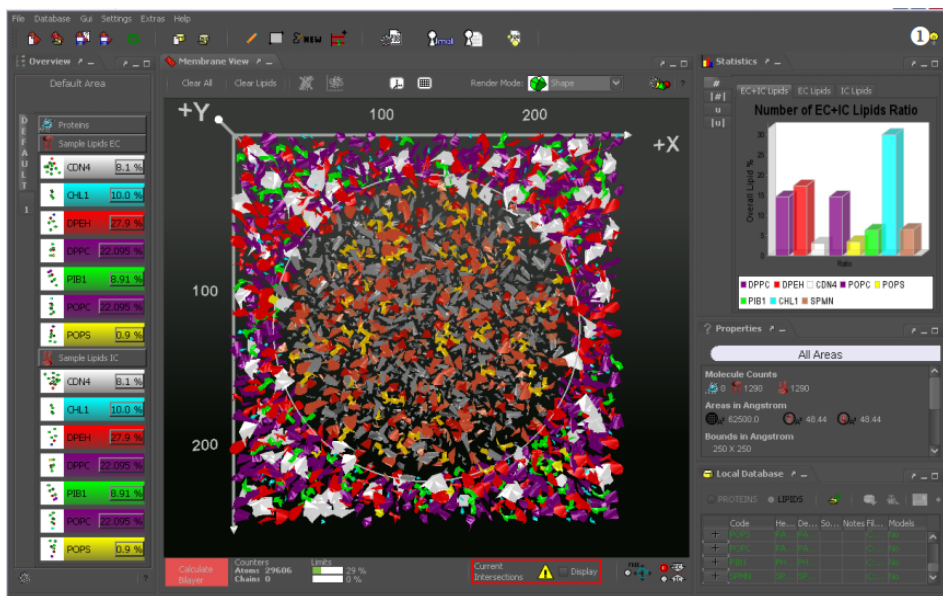


Figura 3.2:

Captura de pantalla del programa CELLmicrocosmos 2.2 MembraneEditor.

Imagen extraída de [Sommer et al., 2011a]

3.5. Armado del modelo de partícula LL5/2

De acuerdo con lo propuesto en los objetivos, se dispuso a validar el modelo LL5/2 propuesto por Silva [Silva et al., 2005] a partir del modelo LL5/5 propuesto por Segrest [Segrest et al., 1999]. Para ello, se partió de la estructura de Segrest, rotando una de las cadenas respecto de la otra alrededor de la circun-

ferencia de la bicapa lipídica hasta lograr la yuxtaposición de los carbonos alfa de las hélices 4 y 8 con los carbonos alfa de las hélices 8 y 10.

Esto se realizó mediante la herramienta *g_confrms*, la cual calcula la desviación cuadrática media (RMSD) después de ajustar por mínimos cuadrados la segunda estructura sobre la primera. Este ajuste se realiza sobre grupos de índices de estructuras, por lo que, en primer lugar fue necesario generar estos índices. En uno de ellos se indicaron los átomos que se utilizaron como estructura inicial y en otro se especificaron los átomos de la estructura a alinear. Para esto, se utilizaron los carbonos alfa de la cadena polipeptídica a fin de eliminar eventuales artefactos debido a los grupos R de los aminoácidos. Los índices correspondientes, fueron realizados con la herramienta *make_ndx* del paquete GROMACS. Se seleccionaron los carbonos alfa de las cadenas 2, 5 y 8 para la estructura de inicio, y los carbonos alfa de las cadenas 5, 8 y 1 de la estructura a alinear, realizándose asimismo los correspondientes índices. Luego se realizó la alineación con *g_confrms*. Dicho esquema se repitió probando varias combinaciones diferentes de cadenas para el alineamiento, sin lograr candidatos exitosos. A partir de la revisión de varios datos experimentales [Silva et al., 2005] se exploró la hipótesis de que las cadenas pudiesen estar corridas una respecto a la otra, aunque sin una superposición exacta sino admitiendo un pequeño desfase. Con esta información se realizaron nuevos índices, correspondientes a los carbonos alfa 195-226, 40-182, 96-118 y 77-140. Nuevamente, dicho esquema se repitió probando varias combinaciones diferentes de cadenas para el alineamiento. Finalmente, seleccionando 3 de estos grupos, correspondientes a los carbonos alfa de los intervalos de átomos 70-89, 133-152 y 188-207, se logró un candidato estructural con el giro deseado.

A continuación se agregó la otra cadena peptídica, que había sido removida, a fin de reconstituir el dímero proteico. Como resultado se obtuvo la estructura correspondiente al modelo buscado LL5/2 (figuras 3.3 y 3.4). Luego se procedió a relajar la estructura proteínica, ya que debido a la rotación de cadenas, algunos residuos se encontraban en posiciones muy cercanas y/o sus cadenas laterales estaban superpuestas. A fin de resolver estas tensiones estructurales se relajó el dímero mediante minimización de energía (EM), usando el algoritmo de *Steepest descent*. Para ello, se convirtió la estructura .pdb, en una estructura compatible con GROMACS a través del programa *pdb2gmx*, que genera un archivo de estructura (.gro) junto con los correspondientes archivos de topología (.top), a partir del campo de fuerza seleccionado, en este caso, el

GROMOS96 43A1.

Adicionalmente se incluyeron restricciones de movimiento para los carbonos alfa de las cadenas proteínicas, ya que las cadenas laterales son las que se encuentran en colisión con otros átomos, o están muy cerca. Para ello, se utilizó la herramienta *genrest* para cada una de las cadenas polipeptídicas por separado. Esta generó un archivo por cadena, conteniendo las restricciones de movimiento de los carbonos alfa. Para que estas restricciones sean consideradas por el programa al momento de minimizar la energía, fueron incluidas manualmente en el archivo de topología.

Se generó el archivo binario de instrucciones de corrida (*.tpr*) con la herramienta *grompp* a partir de un archivo de ingreso de parámetros (*.mdp*), donde se incluye el algoritmo *steep* junto a los archivos de estructura y topología generados. Con este archivo *.tpr* se pudo dar inicio a la corrida con el comando *mdrun*. Con esta corrida se logro un mínimo de energía en menos de 1000 pasos de integración.

A continuación, para relajar aún más la estructura y verificar que no encontremos choques atómicos, se realizó una dinámica de 1 picosegundo, con paso de integración de 0,02 femtosegundos (0.00002 ps), baño térmico *V-rescale* a 5 K y una constante de tiempo *tau_t* de 0,00001 ps, sin baño de presión y con las restricciones de movimiento sobre los carbonos alfa mencionadas anterior-

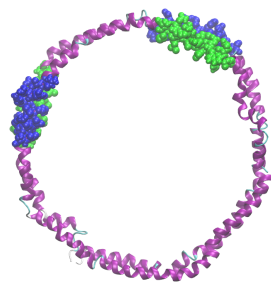


Figura 3.3:
Modelo de cinturón doble en registro LL5/2 generado. En azul se muestran las α -hélices 5 y en verde las α -hélices 2.
Vista superior.

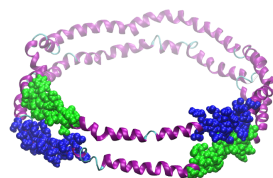


Figura 3.4:
Modelo de cinturón doble en registro LL5/2 generado. En azul se muestran las α -hélices 5 y en verde las α -hélices 2.
Vista lateral.

mente. Para ello, con los datos de salida de la EM, es decir, el archivo de estructura *.gro* generado y los archivos de topología *.top*, se procedió con el mismo esquema, cambiando en el archivo de ingreso de parámetros *.mdp* el algoritmo utilizado, que será ahora *md* (el algoritmo integrador “de a saltos”) [Lindahl et al.,]. De esta forma se obtuvo la estructura de las 2 cadenas de ApoAI en conformación LL5/2.

3.6. Reconstrucción del modelo LL5/5

Se obtuvo la estructura con la cual Segrest realizó la postulación del modelo de cinturón doble en configuración LL5/5, la cual consistía en el dímero de ApoAI en un registro de alfa hélices 5 una sobre la otra, ambas orientadas hacia la izquierda, junto a una bicapa compuesta de 160 moléculas de POPC [Segrest et al., 1999]. Siguiendo el esquema anterior se logró reproducir el modelo LL5/5.

En primer lugar se generó manualmente un archivo de estructura únicamente conteniendo el dímero de ApoAI en configuración 5/5 (figuras 3.5 y 3.6). Se realizó una minimización de energía (EM), siguiendo el algoritmo de *Steepest descent*, de acuerdo al mismo procedimiento para el modelo anterior. Se utilizó herramienta *pdb2gmx*, la cual genera a partir de una estructura, un archivo de estructura *.gro* así como también los correspondientes archivos de topología *.top*, utilizando el campo de fuerza GROMOS96 43A1. Además, se generaron restricciones de movimiento para los carbonos alfa de ambas cadenas. A través de la herramienta *genrest* para cada una de las cadenas polipeptídicas y luego fueron incluídas y definidas manualmente en el archivo de topología. A partir de un archivo de ingreso de parámetros *.mdp*, se generó el archivo binario de instrucciones de corrida *.tpr* con la herramienta *grompp*, definiendo el

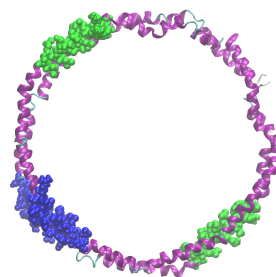
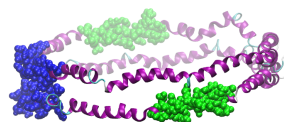


Figura 3.5:
Modelo de cinturón doble en registro LL5/5 reconstruido a partir del modelo de Segrest [Segrest et al., 1999]. En azul se muestran las α -hélices 5 y en verde las α -hélices 2.

Vista superior.

algoritmo *steep* junto a los archivos de estructura y topología generados. Con este archivo *.tpr* se pudo dar inicio a la corrida con el comando *mdrun*. La DM tuvo las siguientes características: duración de 1 picosegundo, paso de integración de 0,02 femtosegundos (0.00002 ps), baño térmico *V-rescale* a 5 K y una constante de tiempo *tau_t* de 0,00001 ps, sin baño de presión y con las restricciones de movimiento sobre los carbonos alfa mencionadas anteriormente [Lindahl et al.,].

Siguiendo el esquema anterior, se generó el archivo binario de instrucciones de corrida *.tpr*, con la herramienta *grompp* bajo el integrador *md* y luego se pudo dar inicio a la simulación con el comando *mdrun*.



3.7. Armado de la bicapa lipídica

De acuerdo al objetivo planteado, se dispuso a generar una bicapa lipídica conteniendo DMPC y Colesterol, en una proporción aproximada de un 90% 1,2-dimiristoil-sn-glicerol-3-fosfatidilcolina (DMPC) y un 10% Colesterol. Para ello, se utilizó el programa CELLmicrocosmos 2.2 MembraneEditor (CmME). Como insumo, este programa necesita archivos de estructura *.pdb* de los monómeros que los integran, en este caso DMPC y Colesterol. Para ello, se obtuvo el archivo *.pdb* de una bicapa de DMPC con colesterol del sitio del Grupo Computational Molecular Biophysics Group, de la Universidad de Gotinga (Georg-August-Universität Göttingen), dirigido por el Dr. Jochen Hub [Wennberg et al., 2012, Hub et al., 2010]. A partir de ésta, manualmente se copiaron los monómeros y se generaron archivos independientes, los cuales se cargaron en la plataforma CmME, y se eligieron sus proporciones, es decir, 90 %

Figura 3.6:
Modelo de cinturón doble en registro LL5/5 reconstruido a partir del modelo de Segrest [Segrest et al., 1999]. En azul se muestran las α -hélices 5 y en verde las α -hélices 2.

Vista lateral.

3.8. COMPATIBILIZACIÓN DEL DÍMERO PROTEICO CON LA BICAPA LIPÍDICA23

DMPC, y 10 % Colesterol. Se partió de una plantilla en blanco de un área de 150 Ångström (Å) cuadrados, y se seleccionó el algoritmo *Distributor*. Este algoritmo utiliza la repulsión entre los lípidos para ordenarlos geoméricamente en una bicapa. Este es el algoritmo indicado para un estado inicial de simulaciones de MD.

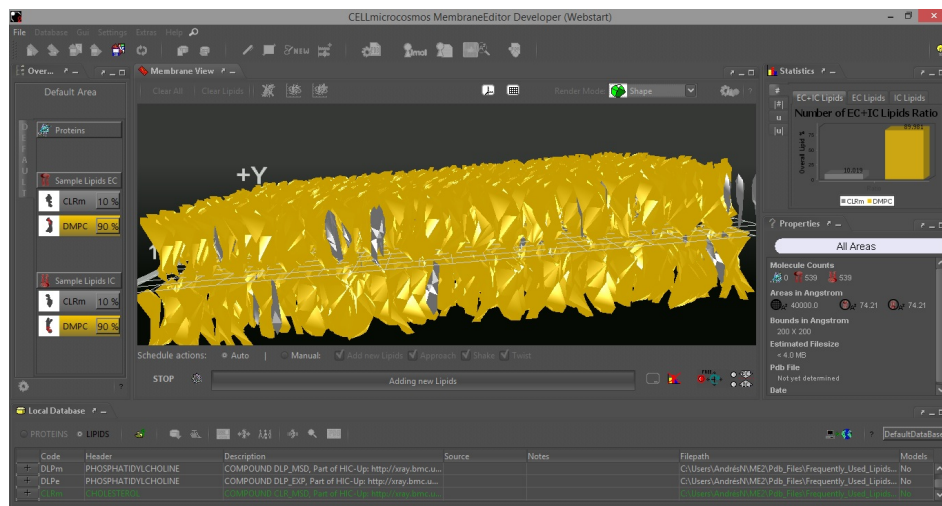


Figura 3.7:

Captura de pantalla del programa CELLmicrocosmos 2.2 MembraneEditor (CmME) mostrando la bicapa generada.

A partir de la bicapa generada, se exportó como un archivo de estructura *.pdb* compatible con DM.

3.8. Compatibilización del dímero proteico con la bicapa lipídica

Se realizó la construcción de los dímeros proteicos en las dos conformaciones deseadas, es decir, LL5/5 y LL5/2, y por otro lado, la bicapa lipídica deseada con 90 % DMPC y 10 % Colesterol. Para generar el modelo completo, se realizó el siguiente armado molecular. En primer lugar, se tomaron los archivos de estructura del dímero proteico y de la bicapa, y se los centró y asignó un tamaño

de caja en común. Esto se realizó a través de la herramienta *Editconf*, con el parámetro *-center*, el cual centra la estructura en la caja seleccionada, y junto con el parámetro *-princ*, que ordena la estructura respecto al eje x, a fin de darle una orientación común a ambas estructuras. Se seleccionó una caja triclínica de un tamaño de 10 x 10 x 10, suficiente para contener las proteínas y la bicapa. Luego, se concatenaron en un archivo de estructura mediante la herramienta *cat*. A partir de este archivo, se actualizó la cantidad de átomos y editaron las líneas correspondientes a los encabezados y pie de los archivos.

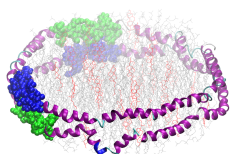


Figura 3.8:

Modelo de cinturón doble en registro LL5/2 con la bicapa lipídica embebida. En azul se muestran las α -hélices 5 y en verde las α -hélices 2.

El archivo concatenado se visualizó con el programa VMD donde se observa claramente la superposición de la bicapa sobre el dímero de ApoAI, verificando que se logró el objetivo deseado.

Con esta herramienta entonces se realizó el siguiente procedimiento para compatibilizar las estructuras. En la pestaña *Graphics* seleccionamos la opción *Representation*, en la cual crearemos una representación deseada con el campo *Selected Atom*. Se seleccionaron a todos los lípidos contenidos en el anillo de proteína cuyos átomos se encuentren a una distancia mayor a 0,8 Ångström (Å) de la misma, es decir, de cualquiera de las dos cadenas

polipeptídicas. Por otro lado, se seleccionaron además, a todos los residuos que se encuentren dentro de una distancia similar al radio de la partícula desde el centro de la misma, a fin de excluir a los lípidos que quedaran por fuera de la circunferencia generada por el dímero de la ApoAI. De esta representación, a través de la herramienta *Save representation as...* en la pestaña *File*, podremos guardar la estructura atómica de los átomos seleccionados. Luego de este procedimiento, el sistema quedó conformado por 2 cadenas polipeptídicas, 238 moléculas de DMPC y 35 de CHOL para el caso del dímero proteico en registro LL5/2 (figura 3.8). Para el caso del dímero proteico en registro LL5/5, el sistema quedó conformado por las 2 cadenas polipeptídicas, 241 moléculas de DMPC y 35 de CHOL (figura 3.9).

Por último, para evitar sobredimensionar la caja, que previamente fue elegida al azar, y que nuestro sistema se encontrara fuera de centro, se volvió a utilizar la herramienta *Editconf* junto a los parámetros: *-c* para centrar todo el sistema en el centro de la caja, *-princ* ordenando todo el sistema respecto al eje x, *-bt triclinic* definiendo la caja triclínica, y *-d 1.8* definiendo el tamaño de la caja en 1,8 nm.

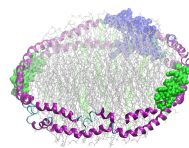


Figura 3.9:
Modelo de cinturón doble en registro LL5/2 con la bicapa lipídica embebida. En azul se muestran las α -hélices 5 y en verde las α -hélices 2.

3.9. Preparación del sistema para las simulaciones

Debido a que el paquete de programas GROMACS, y su herramienta *pdb2gmX* no posee incorporado el campo de fuerza 43A1-S3, se debió modificar manualmente la topología previamente construida únicamente para los dímeros proteicos. Para esto se tomó el archivo *.top* que contiene los datos pertenecientes a la topología del sistema, y se procedió con las ediciones y agregados correspondientes. Así se incluyó dentro de la topología el campo de fuerza 43A1-S3, el cual está preparado para interpretar las estructuras de los lípidos del modelo.

Debido al agregado de la bicapa, se contabilizaron las moléculas agregadas de DMPC y Colesterol a partir del archivo de estructura generado y se agregaron a esta sección correspondiente para la correcta interpretación de la topología. A continuación se realizó una minimización de energía (EM), siguiendo el algoritmo de *Steepest descent*, como fue explicado anteriormente. A partir de esta minimización, la estructura de salida se utilizó para realizar una DM corta, a baja temperatura, paso de integración corto, sin acople de presión y con restricción al movimiento de carbonos alfa, a fin de continuar con la estabilización de la estructura. Para ello se preparó un archivo *.mdp* con los parámetros correspondientes, a partir del cual se generó el archivo binario de instrucciones de corrida *.tpr* con la herramienta *grompp*, definiendo el integrador *md* junto a

los archivos de estructura y topología generados. Con este archivo *.tpr* se pudo dar inicio a la corrida con el comando *mdrun*. Con los datos de salida de esta DM corta, se realizó una nueva DM corta, aumentando el paso de integración para aproximarse a los 2 femtosegundos de una dinámica de producción, y el tiempo de refresco del baño térmico. Con estos datos de salida consideramos que durante 5 ps las cadenas laterales de ambas cadenas de ApoAI, las moléculas de DMPC y colesterol, pudieron recorrer un espacio conformacional lo suficientemente amplio como para lograr acomodarse a su nuevo entorno molecular.

Hasta aquí, nuestro sistema se encontraba en una caja sin solvente. Con la estructura de los dímeros de Apo-A1 relajada y con la bicapa lipídica sin colisiones atómicas, se procedió a la solvatación del sistema. Mediante la herramienta solvate del paquete de programas GROMACS, se agregó un medio acuoso al sistema, eligiendo el modelo de agua SPC/E [Berendsen et al., 1987]. De esta forma se agregaron 48130 moléculas de agua, definiendo al sistema con un total de 160827 átomos. Con este dato se modificó manualmente la topología, quedando así definitivamente definido el sistema.

El procedimiento explicado fue aplicado también para la preparación del sistema correspondiente al modelo de referencia de LL5/5, quedando el sistema definido por 168297 átomos.

3.10. Estabilización

Con el sistema finalmente solvatado, se procedió a la realización de una nueva minimización de energía, esta vez utilizando el archivo de estructura de 168297 átomos. En una primera corrida se utilizó el algoritmo *Steepest descent*, mencionado anteriormente. Luego, repitiendo el esquema anterior, y manteniendo las restricciones de movimiento para el backbone de los dímeros de ApoAI, se procedió a una corrida corta de estabilización, idéntica a la propuesta anteriormente. A continuación se modificó el archivo de ingreso de parámetros *.mdp*, para realizar otra dinámica, reduciendo el paso de integración y el tiempo de acople del baño térmico, y aumentando el paso de integración.

3.11. Adecuación de campo de fuerza a la bicapa elegida

Como ya hemos mencionado, existe un número reducido de estructuras de bicapas lipídicas en archivos PDB disponibles. Este número se reduce aún más cuando se le adicionan los requisitos de ser apta para dinámica molecular y estar estabilizada. Recordemos que el campo de fuerzas utilizado para los lípidos ha sido el 43A1-S3, basado en 43a1 de GROMOS96, y que, como hemos mencionado, la estructura de los monómeros de las moléculas de DMPC y de CHOL corresponden al archivo *.pdb* de una bicapa de DMPC con colesterol. Esto nos forzó a tener que editar el campo de fuerzas, debido a que los nombres de los átomos de una y otra fuente eran distintos. Esta edición se realizó manualmente con cada monómero. En primer lugar se tomó el archivo de estructura correspondiente a la bicapa y manualmente se generó un archivo conteniendo una única molécula de DMPC, que se visualizó con el programa VMD y se realizó un esquema anotando el nombre de cada átomo. Con esta información, se procedió a la edición del archivo *lipids_43A1-S3.itp*, parte del campo de fuerzas 43A1-S3, el cual contiene la información relacionada a la topología de las moléculas lipídicas que contiene el campo de fuerza. A partir de este, se generó una versión propia de este archivo anotando la nomenclatura definida y se renombró los átomos dentro del campo de fuerza con la nueva anotación. Como verificación se simuló la bicapa de partida completa durante 5ns para evaluar su comportamiento y se observó un comportamiento normal, correspondiente al modelo de mosaico fluido.

Capítulo 4

Resultados y discusión

La técnica empleada para el estudio de este sistema es la simulación computacional por Dinámica Molecular. En particular se utiliza el paquete GROMACS (sigla en inglés para GROningen MACHine for Chemical Simulations) 5, 6, 7, 8, herramienta de acceso libre y gratuito, de probada eficiencia en este campo, y ampliamente utilizada en todo el mundo. Dentro de las facilidades de este paquete se encuentran herramientas de análisis como SASA (superficie accesible al solvente), *RMSD* (desviación cuadrática media), *Order parameters* (parámetro de orden), *Distance* (Distancia) y *DSSP* (diccionario de patrones de estructura secundaria) que fueron utilizadas para evaluar los resultados.

Para la realización de la simulación computacional del modelo planteado de la Apolipoproteína A1 humana, se empleó el paquete GROMACS 5.0.6 [Eargle et al., 2006, Frishman and Argos, 1995, Humphrey et al., 1996, Sanner et al., 1995, Sharma et al., 2000, Stone et al., 2001, Stone, 1998, Varshney et al., 1994]. Se utilizó el campo de fuerzas para átomos enlazados 43a1 de GROMOS96 (ffg43a1) 93 [Berger et al., 1997], tanto para los procedimientos de estabilización del sistema, como para todos los pasos de simulación por Dinámica Molecular. Para los lípidos, se utilizó el campo de fuerzas 43A1-S3 [Chiu et al., 2009], adecuado para bicapas lipídicas. Para todas las simulaciones de este sistema se utilizó un clúster de 32 nodos corriendo bajo GNU/Linux, utilizando para las visualizaciones de trayectorias y estructuras el software de distribución libre Visual Molecular Dynamics (VMD). Para las representaciones gráficas de resultados se utilizó el programa GRACE. Para los estudios en solvente acuoso se utilizó el

modelo de agua SPC/E [Berendsen et al., 1987]. Las interacciones de Lennard-Jones fueron calculadas explícitamente dentro de un radio de corte de 1,0 nm. Se mantuvieron las restricciones topológicas para las longitudes de enlace utilizando el algoritmo LINCS. Para este sistema se utilizó una caja de simulación triclinica.

4.1. Detalles de las simulaciones

A continuación se realizaron dos simulaciones consecutivas de 5 ns cada una, dando un total de 10 ns. Los análisis presentados se realizaron con los segundos 5 ns de simulación.

En las simulaciones de 5 ns la presión fue mantenida constante acoplando isotrópicamente el sistema con el baño de presión de *Berendsen* de un 1 bar, con una constante de acoplamiento (τ_p) de 2 ps y una compresibilidad de $4.5 \times 10^{-5} \text{ bar}^{-1}$.

La temperatura se constante, con una constante de acoplamiento (τ_t) de 0,1 ps, acoplando las proteínas, los lípidos y el solvente por separado, a una temperatura de baño de 300K usando el algoritmo *V-rescale*.

Todas las simulaciones se realizaron con condiciones periódicas de contorno. El paso de integración fue de 2 fs en todos los casos.

4.2. Análisis de las simulaciones

Todos los análisis se realizaron para los 2 modelos planteados, es decir el LL5/5 y el LL5/2, a los últimos 5 ns de simulación.

4.2.1. RMSD

La sigla *RMSD* proviene del inglés de desviación cuadrática media (root-mean-square deviation) y es la medida de la distancia promedio entre átomos. Se utiliza para el estudio de distintas moléculas, principalmente proteínas, y es una forma de comparar las variaciones en las estructuras iniciales y finales. Este parámetro puede ser calculado comparando las estructuras obtenidas a cada tiempo contra todo el tiempo de la simulación, siendo de interés ya que nos da una idea de que conformaciones recorre nuestra proteína. Con la ayuda de la herramienta *gmx rms* del paquete Gromacs, se le dió como insumo la

estructura de referencia, es decir, la estructura de inicio de los últimos 5 nanosegundos de simulación, junto a la trayectoria de los átomos en los 5 nanosegundos de dinámica. Para el cálculo, se seleccionó el grupo de átomos correspondientes a los carbonos alfa de la estructura, para monitorear los cambios en distancia de la proteína, sin estar afectados ni por las cadenas laterales de la misma ni por el movimiento de los lípidos. A partir de los datos de salida, estos fueron convertidos a una matriz de puntos y luego fueron coloreados para una mejor visualización.

En la matriz graficada para el modelo LL5/5 (figura 4.1) se identifica la diagonal identidad, propia de este tipo de gráficos, y las estructuras adyacentes inmediatas con colores azules y celestes, correspondientes a valores bajos de RMSD, ya que se comparan cuadros muy cercanos en el tiempo. Se observa una predominancia del color verde, correspondiente a valores de RMSD de intermedios, de 1.5 Å y 2.5 Å. Los colores que corresponden a los valores de RMSD más altos, que se observan en rojo, se encuentran en baja proporción, y especialmente concentrados cerca de los cuadros correspondientes a 4 nanosegundos de simulación, para luego retornar a valores intermedios. A partir de este gráfico podemos decir que el modelo LL5/5 recorre un espacio conformacional propio, manteniendo una fuerte similitud con su estructura de partida, sin la formación de conformeros distintos.

Para el modelo LL5/2 (figura 4.2), se identifica también la diagonal identidad y los valores bajos cercanos a la misma. Se observa que a medida que avanza la simulación se encuentran valores de RMSD que aumentan progresivamente, pero manteniéndose en valores comparables a los obtenidos en la estructura LL5/5. En este caso tampoco observamos formación de clústers de estructuras. A partir de este gráfico podemos decir entonces, que el modelo LL5/2 recorre su espacio conformacional, pero manteniendo una fuerte similitud con la estructura de partida. En comparación con el modelo LL5/5, se observa una fuerte similitud en ambos modelos, y que ambos obtienen valores máximos de RMSD similares. Por otro lado, el tiempo de simulación es acotado, por lo que para profundizar este análisis y buscar posibles conformeros distintos se requeriría un mayor tiempo de simulación.

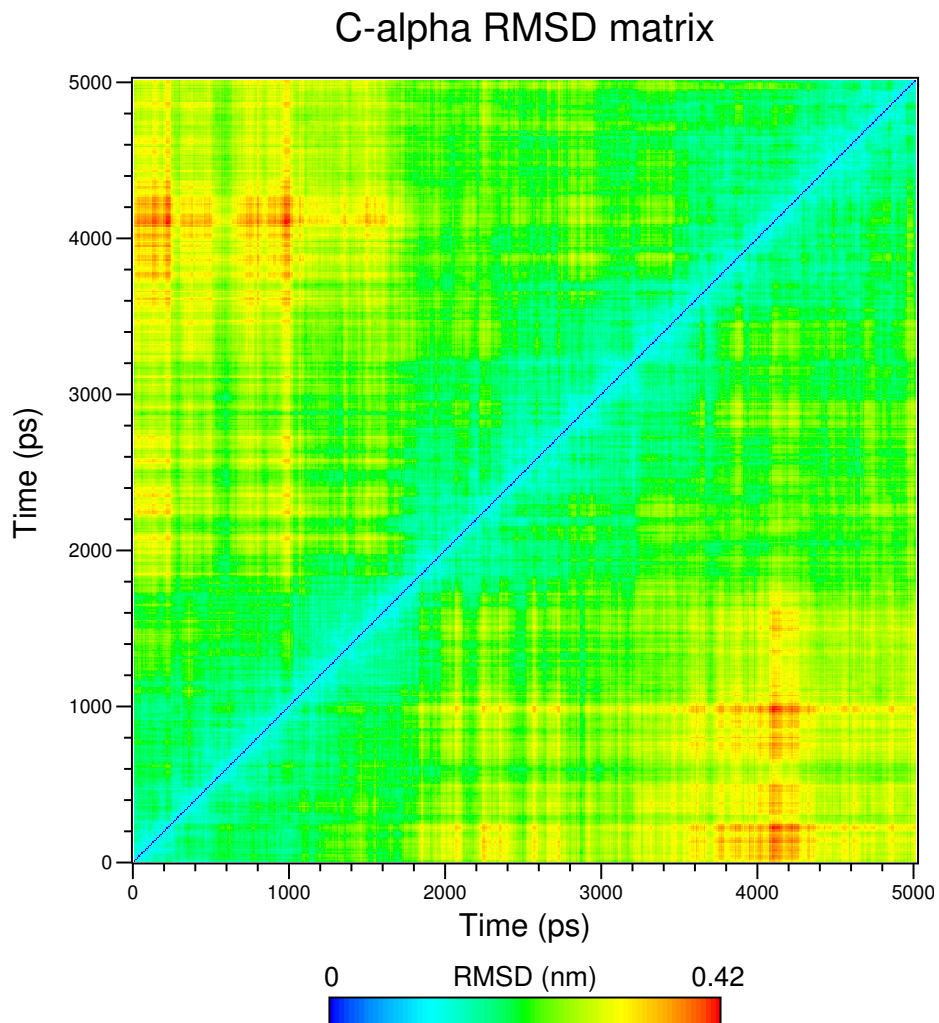


Figura 4.1:
Matriz de puntos generada a partir de el análisis de RMSD para el modelo LL5/2. En esta se compara cada cuadro contra todos los cuadros de la simulación y se gráfica con una escala de colores. Tamaño: 500 x 500.

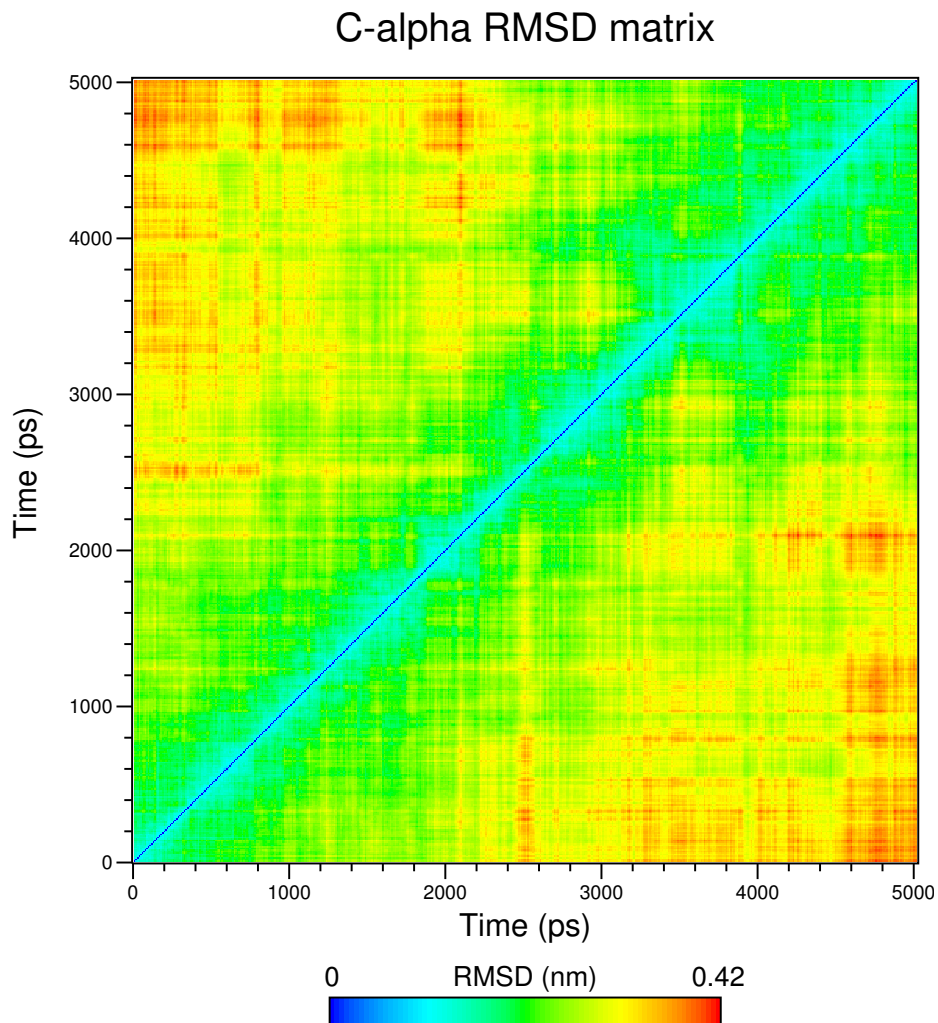


Figura 4.2:
Matriz de puntos generada a partir de el análisis de RMSD para el modelo LL5/2. En esta se compara cada cuadro contra todos los cuadros de la simulación y se gráfica con una escala de colores. Tamaño: 500 x 500.

4.2.2. Análisis de superficie accesible al solvente (SASA)

La herramienta SASA del paquete de programas GROMACS, es una herramienta que nos permite realizar el cómputo de la superficie accesible al solvente. Esta utiliza el algoritmo de doble red cúbica (double cubic lattice) implementado por Eisenhaber et al [Eisenhaber et al., 1995]. La superficie accesible al solvente se refiere a la parte de la superficie la cual, una esfera con centro en un átomo de radio definido, puede ser puesta en contacto con el radio atómico de van der Waals, sin penetrar otros átomos. Es decir, es el volumen molecular de la cual el solvente se encuentra excluido.

En nuestro caso, nuestro modelo excluirá al solvente mediante interacciones de dos grupos distintos. En primer lugar, como cualquier proteína, parte del solvente será excluido por la proteína en sí misma. Por otro lado, la bicapa excluirá a otra parte del solvente.

A fin de obtener una mayor información, se buscó obtener datos por separado del dímero de ApoAI y de los lípidos a lo largo de toda la simulación. Para obtener estos datos, se debió definir las salidas de datos del programa, modificando el mismo para que lograra discriminar entre ambos. Para ello se modificó el comando *gmx sasa* definiendo la superficie a calcular con el parámetro *-surface* y definiendo los grupos de salida con el parámetro *-output*. De esta forma se logró que calculara por separado un valor de SASA correspondiente a la interacción proteica con el solvente y otro correspondiente para la interacción lipídica con el solvente. Estos valores de SASA fueron calculados durante los 5 ns de simulación. Los datos obtenidos fueron graficados con el graficador *Grace*.

Los gráficos presentan en rojo la porción correspondiente al valor de superficie accesible al solvente de los dímeros proteicos, y el valor en verde corresponde al conjunto de la bicapa lipídica, es decir, DMPC y Colesterol. Además, se presenta el valor total del sistema, correspondiente a la suma algebraica de los dos anteriores para cada punto.

Para el modelo LL5/5 (Figura 4.3) , se observan que ambas curvas oscilan entre un valor definido y se mantienen durante toda la simulación. Además, podemos observar que la porción proteica presenta una desviación de su valor promedio muy baja. Eso explica la baja movilidad de la porción proteica, y denota la estabilidad del modelo. En cambio la porción lipídica, presenta una mayor variación respecto de su valor promedio, comportamiento esperable y

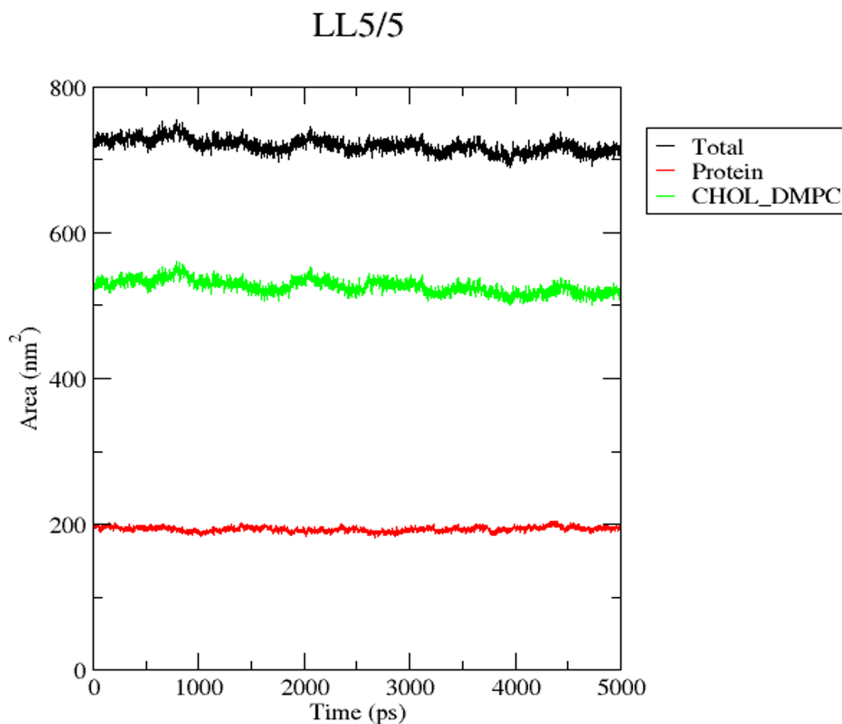


Figura 4.3:
Gráfico de área accesible al solvente durante el tiempo de simulación para el modelo LL5/5. En rojo vemos la fracción proteica. En verde la fracción lipídica. En negro la correspondiente al total del sistema.

comparable a otros trabajos [Berger et al., 1997].

Para el modelo LL5/2 (Figura 4.4), se observa un comportamiento similar al presentado por la simulación LL5/5. También se observa que la porción proteica presenta una desviación de su valor promedio muy baja, indicando que la estabilidad del modelo LL5/2 es comparable a la del LL5/5. La porción lipídica presenta una comportamiento comparable al modelo LL5/5.

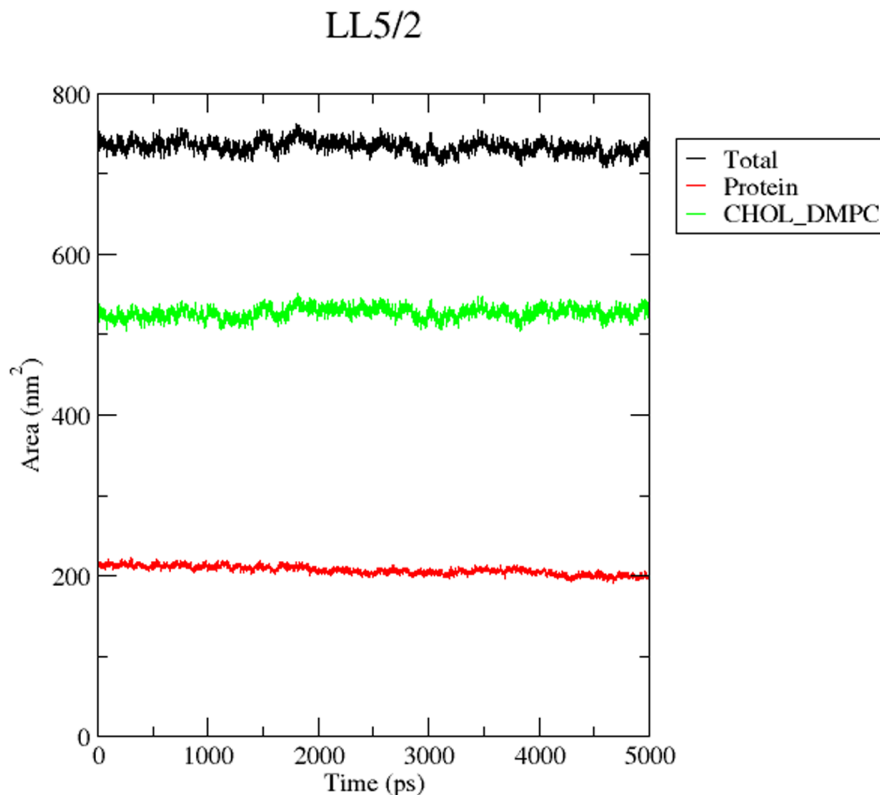


Figura 4.4:
Gráfico de área accesible al solvente durante el tiempo de simulación para el modelo LL5/2. En rojo vemos la fracción proteica. En verde la fracción lipídica. En negro la correspondiente al total del sistema.

4.2.3. Análisis de estructura secundaria

Para realizar este análisis, se utilizó el programa *DSSP*, sigla en inglés para "diccionario de patrones de estructura secundaria". Este programa fue creado para estandarizar la asignación de estructura secundaria de proteínas. Analiza e identifica la estructura secundaria dada a la estructura tridimensional de

Secondary structure

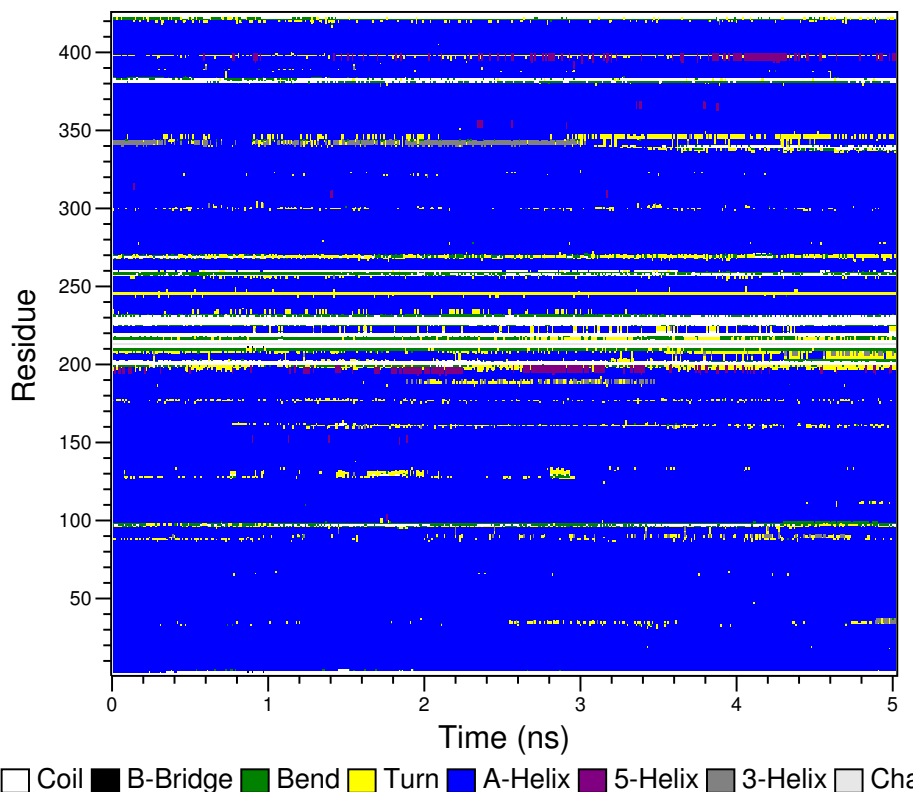


Figura 4.5:

Matrix de salida del programa DSSP de la trayectoria de la simulación del modelo LL5/5. El color blanco corresponde a una estructura tipo random coil.

El color azul corresponde a una estructura de α -helice. El color bordo corresponde a una estructura tipo 5 helice (π -helice). El color amarillo corresponde a turn (giro). El color verde corresponde a bend (dobles).

la proteína y también en la trayectoria de la dinámica molecular de la misma. Es decir, realiza este análisis a través del tiempo y vuelca esta información en un gráfico de dos dimensiones, tiempo vs. residuo. En este gráfico se obser-

van las distintas estructuras secundarias de cada residuo a través del tiempo, y diferenciados con colores característicos (Kabsch & Sander, 1983; Touw et al., 2015).

Secondary structure

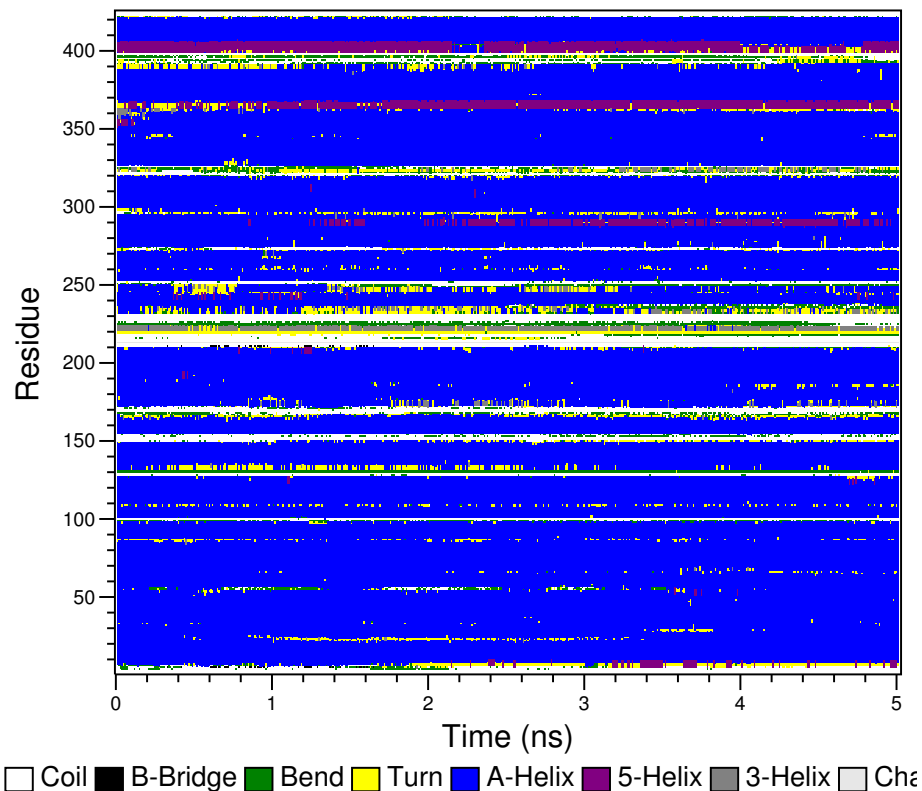


Figura 4.6:

Se observa la matrix de salida del programa DSSP de la trayectoria de la simulación del modelo LL5/2. El color blanco corresponde a una estructura tipo random coil. El color azul corresponde a una estructura de α -helice. El color bordo corresponde a una estructura tipo 5 helice (π -helice). El color amarillo corresponde a turn (giro). El color verde corresponde a bend (dobles).

Este programa, aunque incluido en el paquete Gromacs, hay que descargarlo e instalarlo independientemente. Luego, se deberá definir en las variables del entorno para que el paquete Gromacs lo reconozca. Para una mejor visualización de los resultados se construyeron matrices con una trayectoria reducida.

En la figura 4.5 se observa un predominio mayoritario de α -hélices. Esto es esperable y corresponde con todos los datos en bibliografía. También con lo observado durante el armado del modelo y tratamiento de las estructuras. Esta estructura secundaria predominante se mantiene en el tiempo, dando cuenta de la estabilidad del modelo. Se observa también, que a lo largo de la simulación, se ve como distintas regiones de la estructura recorren distintas conformaciones en un equilibrio dinámico.

En la figura 4.6 se observa también un predominio mayoritario de estructura de α -hélice y como las distintas regiones de la estructura recorren algunas conformaciones, manteniéndose en un equilibrio dinámico. Sin embargo, no se presentan cambios irreversibles en las estructuras secundarias durante el intervalo de la simulación, manteniendo la estructura de alfa hélice. En comparación con el modelo LL5/5 podemos observar también el predominio de α -helice, y que existen más regiones de Turn o hélice 5. Estas estructuras secundarias son más relajadas que al de α -helice, y tienen una mayor dinámica, y esto nos da una idea que el modelo LL5/2 se encuentra más relajado.

4.2.4. Cálculos de deuterium order parameters

Cuando se simulan moléculas con largas cadenas carbonadas, como es el caso las moléculas de DMPC de nuestro modelo, es de interés poder calcular su orientación promedio. Para ello, se utiliza como medida, la desviación de las cadenas carbonadas respecto de un vector perpendicular a la membrana. Esta es la forma por la cual el order parameter mide el comportamiento de las cadenas carbonadas. Existen varias formas de calcularlo, el método más utilizado es el determinado experimentalmente a través de espectrometría resonancia magnética nuclear de deuterio (*deuterium NMR*). Se eligió medir este parámetro independientemente para cada cadena carbonada. Esto se debe a que muchas veces, por la geometría diferencial de ambas cadenas, se presentan comportamientos levemente distintos, y con la unificación de estos datos se perdería esa información. En este caso, la diferencia sustantiva se refiere a la

geometría de la molécula, por la cual una cadena se encuentra levemente más cerca de la cabeza polar de la misma.

Este parámetro entonces fue calculado de la siguiente forma. En primer lugar y ya que durante la simulación, la partícula completa se fue trasladando y rotando dentro la caja, se debió alinear la misma con el eje z. Para ello, se tomaron los carbonos alfas de la proteína y con estos se ordenó la partícula completa respecto del eje Z, con la herramienta *gmx trjconv*, con el parámetro *-fit rot+trans* que realiza un ajuste de la trayectoria respecto de una estructura de referencia, es decir, una en la cual nuestra partícula se encuentre perpendicular al eje z.

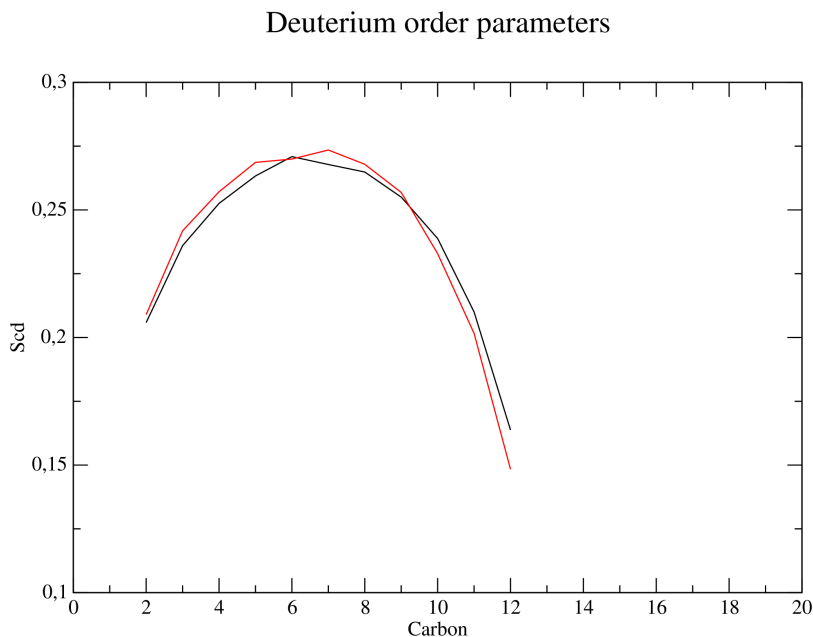


Figura 4.7:
Calculo de deuterium order parameters para el modelo LL5/5. Cada una de las 2 curvas representa una de las colas hidrocarbonadas de las moléculas de DMPC.

Por otro lado, fue necesario crear un archivo index, para indicar cuáles son los átomos correspondientes a las cadenas alifáticas de las moléculas de DMPC. Para ello, se anotaron los nombres de los átomos de las cadenas alifáticas mediante el programa de visualización VMD, y a continuación, con la herramienta *make_ndx* se realizaron los 2 archivos necesarios.

Por último, con la herramienta *gmx order*, que realizó el cálculo de order parameters, usando como insumo el archivo de trayectoria mencionado anteriormente, la estructura de referencia, el archivo index, y además un archivo de la topología del sistema. Los datos de salida fueron graficados y visualizados con el programa *Grace*.

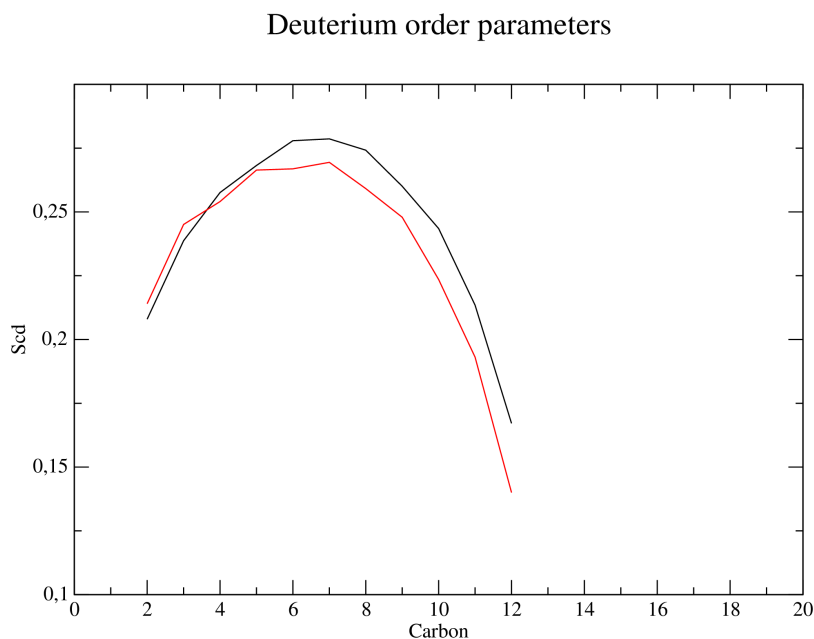


Figura 4.8:
Cálculo de deuterium order parameters para el modelo LL5/2. Cada una de las 2 curvas representa una de las colas hidrocarbonadas de las moléculas de DMPC.

A partir de la figura 4.7 se observan dos curvas correspondientes a las distintas cadenas alifáticas de las moléculas de DMPC. Entre ellas, prácticamente no encontramos diferencias, por lo que podemos afirmar que ambas cadenas tienen un comportamiento similar.

También se construyó de la misma forma el gráfico para la trayectoria del modelo LL5/2. A partir de la figura 4.8 se observa al igual que el anterior dos curvas, correspondientes a las dos cadenas alifáticas de las distintas moléculas de DMPC. Tampoco observamos una diferencia entre ellas, por lo que podemos afirmar que ambas cadenas tienen un comportamiento similar. Por otro lado podemos afirmar que el comportamiento de los lípidos en ambos modelos es comparable. Vale aclarar que ambos modelos muestran a los lípidos de membrana con un orden mucho mayor al correspondiente a una bicapa lipídica libre [Chiu et al., 2009].

4.2.5. Distancias entre residuos

Una vía para validar los modelos planteados radica en observar las distancias entre residuos específicos, ya sea entre residuos intracatenarios o intercatenarios. Para referenciar nuestro trabajo, se utilizó el trabajo de Silva y otros [Silva et al., 2005], quienes a partir de aislados de ApoAI de plasma humano, reconstituyeron las HDL nacientes (rHDL). A estas rHDL se les agregó Bi sulfo-succinimidil suberato (BS3), que es capaz de unir los átomos de nitrógeno de la cadena lateral de las lisinas. Estos enlaces, se forman únicamente si la distancia entre estos átomos es menor o igual a 11,4 Ångström. Esto se debe a que este es el tamaño del BS3. Los resultados obtenidos de estos ensayos fueron evaluados posteriormente por espectrometría de masas, con el fin de determinar cuáles enlaces era posible que se formaran en las rHDL formadas.

Modelo LL5/5

Además, en el trabajo se presenta esquemáticamente los enlaces en cada modelo. En este caso se seleccionó el modelo de cinturón doble, en un registro de LL5/5.

A partir de estos enlaces compatibles con el modelo LL5/5, se planteó analizarlos y compararlos con la simulación obtenida de este modelo. Para ello, con la herramienta VMD, se visualizó la estructura de los polipéptidos, se identificó a cada lisina involucrada en ese trabajo, y se determinó el número de áto-

Lisinas	Modelo LL5/5
K208-K208	Incompatible
K88-K118	Incompatible
K96-K118	Incompatible
K40-K239	Compatible
K118-K40	Compatible
K59-K208	Compatible
K226-K239	Compatible
K59-K195	Compatible
K77-K195	Compatible

Cuadro 4.1:

Tabla de distintos entrecruzamientos encontrados por Silva [Silva et al., 2005] y su compatibilidad con el modelo LL5/5.

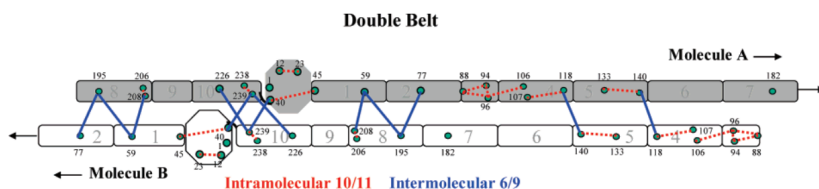


Figura 4.9:

Representación esquemática de los enlaces generados por entrecruzamiento encontrados experimentalmente en el modelo LL5/5. Una cadena de Apo-A1 se muestra en gris y la otra en negro. Los entrecruzamientos intracatenarios se muestran en rojo, y los intercatenarios en azul.

Figura extraída de [Silva et al., 2005]

mo correspondiente para el Nitrógeno de la cadena lateral de dicha Lisina. Los enlaces utilizados figuran a continuación:

Con esta información, se generó un índice a través de la herramienta *make_ndx*, y luego se agregó un apartado de *[Bonds]* conteniendo los átomos a utilizar. A partir de la herramienta *gmx distance* se obtuvieron una serie de gráficos, siendo de nuestro interés el parámetro *-oall*, correspondiente a todas

Cadena peptídica A		Cadena peptídica B	
Lisina	Átomo número	Lisina	Átomo número
K40	70	K239	4335
K239	2144	K40	2261
K118	928	K140	3345
K140	1154	K118	3119
K59	272	K208	4023
K208	1832	K59	2463
K226	2006	K239	4322
K239	2144	K226	4197
K59	272	K195	3901
K195	1710	K59	2463
K77	2661	K195	3901
K195	1710	K77	470

Cuadro 4.2: Tabla de conversión para el modelo LL5/5, relacionando cada lisina de cada cadena peptídica con su número de átomo de nitrógeno

las distancias individuales. Luego, con el programa graficador Grace, se analizaron los datos obtenidos mediante *Running Average*, con un promedio de 15 puntos, es decir 15 ps. Los gráficos se realizaron de a pares, seleccionando los enlaces opuestos para una mejor observación. En todos se gráfico una constante en 1.14 nm, valor de interés ya que a igual o menor distancia, se podrían formar los enlaces encontrados en los datos bibliográficos.

En la figura 4.10 se observan las distancias correspondientes a las lisinas opuestas 40-239. Ambas distancias presentan un comportamiento similar, y oscilan en el tiempo, comportamiento esperable para una simulación de DM. Este par de enlaces, se encuentran lejos del valor de 1.14 nm para toda la simulación, por lo que podemos incluir que estos enlaces no se formarían bajo la acción del agente BS3.

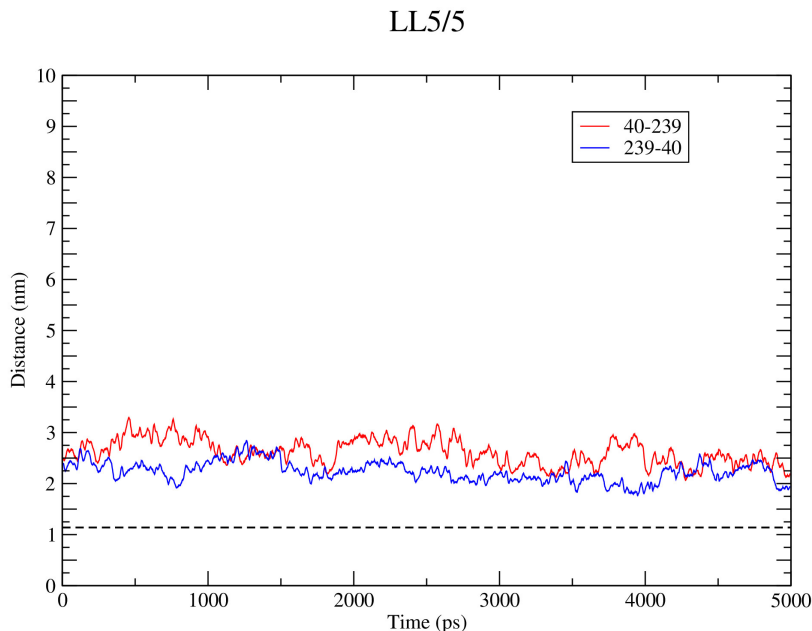


Figura 4.10:

AAAnálisis de distancias en los 5 ns de simulación del modelo LL5/5 los N de las lisinas 40-239 y 239-40.

En la figura 4.11 Se observan el par de distancias de K59-K195 y K195-K59. Comparable al gráfico anterior, ambas presentan un comportamiento similar, oscilando entre 1,25 nm y 2,75 nm. Durante toda esta simulación, ambos enlaces se encontraron a distancias superiores a 1,14 nm y las reacciones con BS3 no serían observadas.

En la figura 4.12 se observa la distancia del par de lisinas 59-208, con comportamiento similar a las presentadas anteriormente. En este caso, ambas distancias se encuentran en algún momento de la simulación a una distancia igual o menor a 1,14 nm, por lo que podemos concluir que, el entrecruzamiento podría ser observado en la simulación realizada.

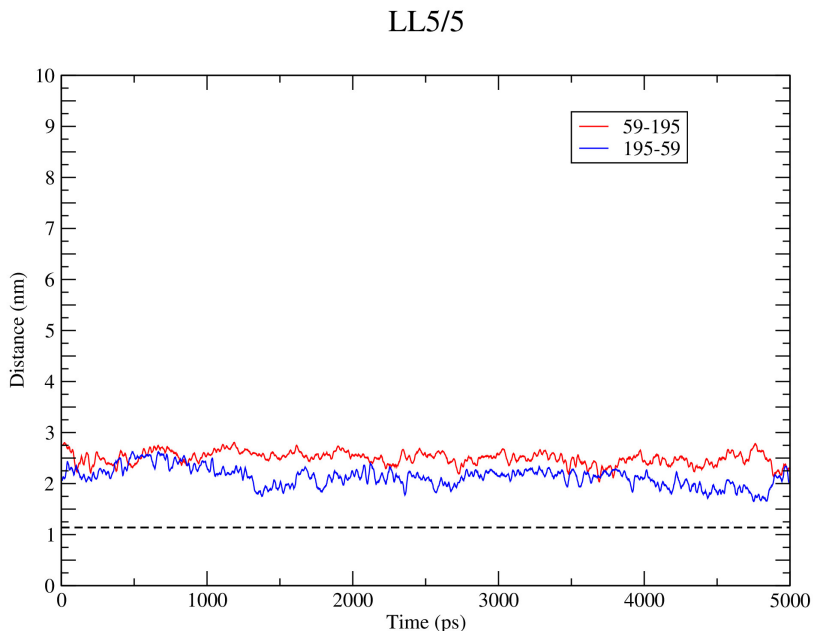


Figura 4.11:
Análisis de distancias en los 5 ns de simulación del modelo LL5/5 entre los N
de las lisinas 59-195 y 195-59.

En la figura 4.13 se observa la distancia del par de lisinas opuestas 77-195. Ambas distancias poseen comportamientos similares y oscilantes, como las presentadas anteriormente. Estas presentan una gran cantidad de tiempo de residencia durante la simulación a una distancia igual o menor a 1.14 nm. En nuestro modelo LL5/5, el entrecruzamiento observado por el agente BS3 para el enlace 77-195, se formaría en ambos pares de lisinas.

En la figura 4.14 se observa la distancia del par de lisinas opuestas 118-140. Se observan ambas distancias con comportamientos similares y oscilantes, como las presentadas anteriormente. Se observa que ambos presentan algún tiempo de residencia por debajo del valor de 1.14 nm. Esto nos dice que en

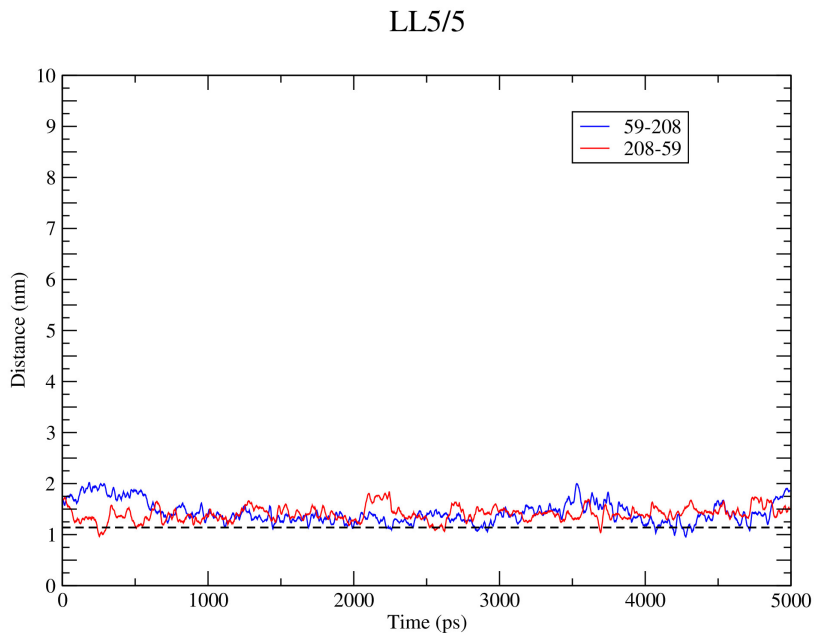


Figura 4.12:
Análisis de distancias en los 5 ns de simulación del modelo LL5/5 entre los N
de las lisinas 59-208 y 208-59.

nuestro modelo LL5/5, el entrecruzamiento observado por el agente BS3 en el trabajo de referencia se formaría.

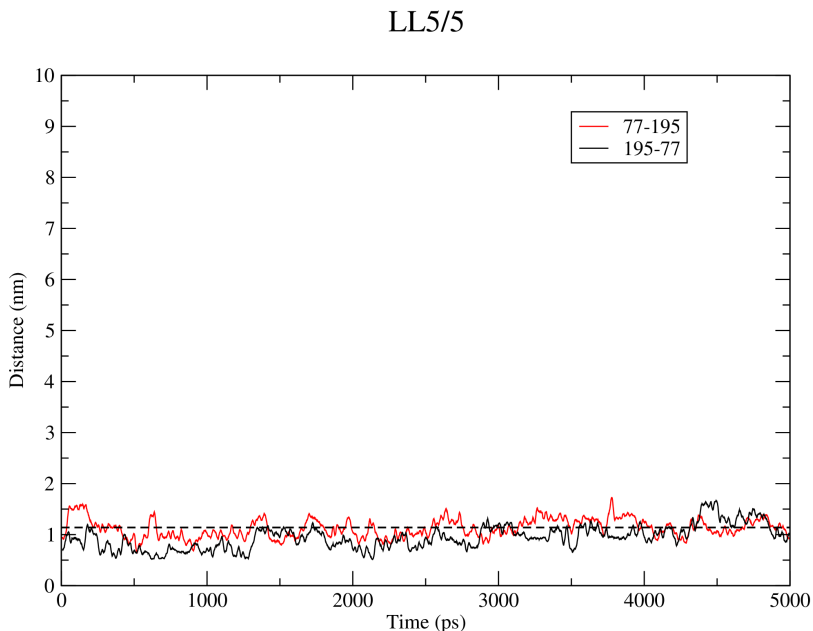


Figura 4.13:
Análisis de distancias en los 5 ns de simulación del modelo LL5/5 entre los N
de las lisinas 77-195 y 195-77.

En la figura 4.15 se observa la distancia del par de lisinas opuestas 226-239. Se observan ambas distancias con comportamientos similares y oscilantes, como las presentadas anteriormente. Se destaca la diferencia del valor del promedio de cada distancia, dando una idea de que un par se encuentra considerablemente más cerca que su opuesto. Aun así, ninguno presenta tiempo de residencia por debajo del valor de 1.14 nm. Esto nos dice que en nuestro modelo LL5/5, el entrecruzamiento observado por el agente BS3 en el trabajo de referencia no se formaría.

De las distancias medidas y los enlaces propuestos, la dinámica del modelo LL5/5 planteado pudo satisfacer 3 de los 6 enlaces de los experimentos

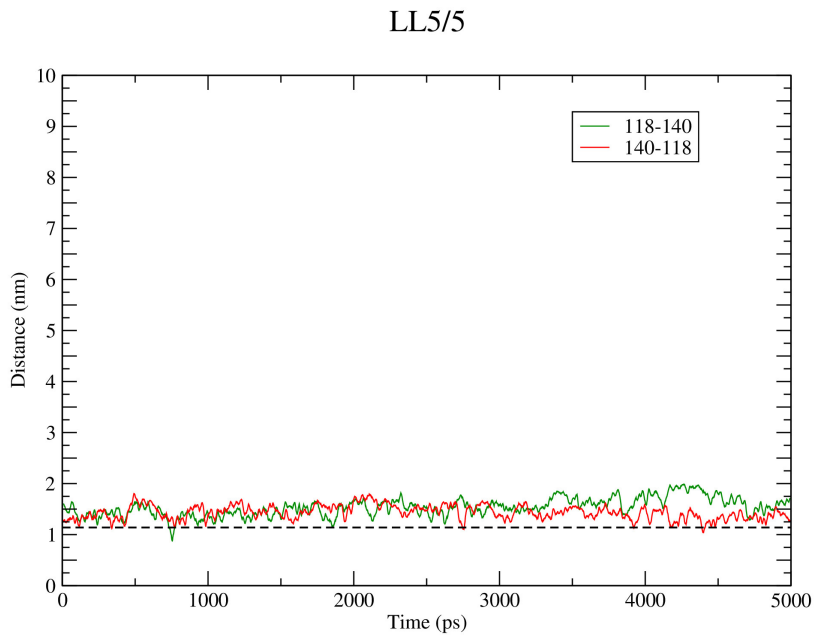


Figura 4.14:
Análisis de distancias en los 5 ns de simulación del modelo LL5/5 entre los N
de las lisinas 118-140 y 140-118.

planteados por Silva.

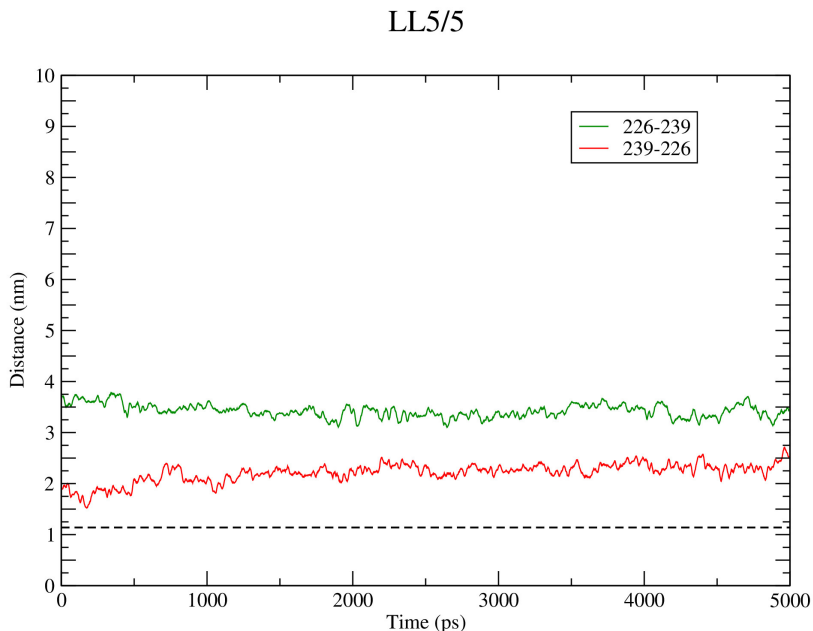


Figura 4.15:
Análisis de distancias en los 5 ns de simulación del modelo LL5/5 entre los N
de las lisinas 226-239 y 239-226.

Modelo LL5/2

En el trabajo de Silva, también se presentaron enlaces que no podían ser explicados por el modelo del cinturón doble en registro de LL5/5, por lo cual, se presenta el modelo LL5/2 como una alternativa posible, para explicar dichos enlaces. Tabla 4.3.

Además, similar al modelo LL5/5, se presenta esquemáticamente los enlaces en un registro de LL5/2 en la figura 4.16.

A partir de estos datos de enlaces por entrecruzamiento para modelo LL5/2, se realizó el mismo procedimiento, utilizando estas lisinas propuestas. Los enlaces propuestos, se identificó a cada lisina involucrada, y se ubicó el número

Residuos	Modelo LL5/2
K208-K208	Posible
K88-K118	Posible
K96-K118	Posible
K106-K106	Posible

Cuadro 4.3: Tabla de distintos entrecruzamientos "huerfanos" del modelo LL5/5 encontrados por Silva [Silva et al., 2005] y su compatibilidad con el modelo LL5/2.

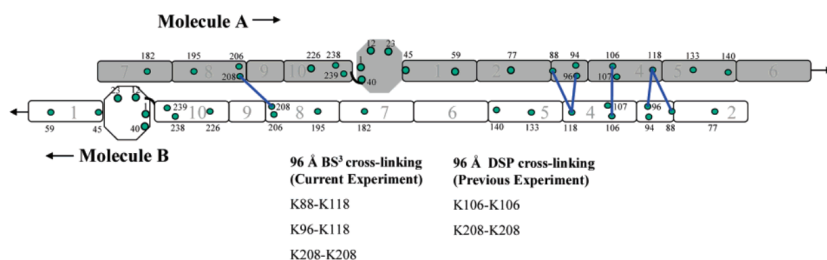


Figura 4.16:

Representación esquemática de los enlaces generados por entrecruzamiento encontrados experimentalmente en el modelo LL5/2. Una cadena de Apo-A1 se muestra en gris y la otra en negro. Los entrecruzamientos intercatenarios se muestran en azul. Figura extraída de [Silva et al., 2005]

de átomo correspondiente para el Nitrógeno de dicha Lisina. Los enlaces se encuentran en la tabla 4.4.

Al igual que para el modelo LL5/5, se obtuvieron los datos correspondientes a todas las distancias individuales. Luego, con el programa Grace, se analizaron los datos obtenidos mediante *Running Average* con un promedio de 15 puntos, es decir 15 ps. Los gráficos se realizaron de a pares y se graficó la constante en 1.14 nm.

En la figura 4.17 se observa la distancia del par de lisinas opuestas 88-118. Ambas distancias poseen comportamientos similares y oscilantes y tienen un tiempo de residencia considerable por debajo del valor de 1.14 nm. Por lo que en nuestro modelo LL5/2, el entrecruzamiento observado en el trabajo de re-

Cadena peptídica A		Cadena peptídica B	
Lisina	Átomo número	Lisina	Átomo número
K208	1832	K208	4023
K88	582	K118	3119
K118	928	K88	2773
K96	660	K118	3119
K118	928	K96	2851
K106	774	K106	2965

Cuadro 4.4: Tabla de conversión para el modelo LL5/2, relacionando cada lisina de cada cadena peptídica con su número de átomo de nitrógeno

ferencia se formaría.

En la figura 4.18 se observa la distancia del par de lisinas opuestas 96-118, donde ambas distancias se encuentran por debajo del valor de 1.14 nm la mayor parte del tiempo de simulación. Podemos decir que para este enlace, el entrecruzamiento observado se formaría, validando el modelo LL5/2.

En la figura 4.19 se observan dos enlaces que no son opuestos, sino que son la distancia a la misma lisina del otro péptido. Debido a esto, únicamente encontramos una sola distancia a evaluar. Para la distancia de las lisinas 208, se observa claramente un comportamiento oscilante, similar a los anteriores. Esta distancia se encuentra en valores superiores a los necesarios para la formación del enlace, por lo que este enlace no se formaría en nuestro modelo. Para el enlace de las lisinas 106, se observa claramente un comportamiento oscilante, similar a los anteriores. Esta distancia presenta valores muy cercanos al valor de 1.14 nm, por lo que no podemos afirmar o negar la formación de un enlace bajo el agente BS3.

De las distancias medidas y los enlaces propuestos, la dinámica del modelo LL5/2 planteado pudo satisfacer 2 de los 4 enlaces de los experimentos planteados por Silva. Además, el enlace entre lisinas 106, se encuentra muy cerca del umbral. Para poder validar definitivamente un modelo u otro, se deberían continuar las simulaciones a fin de monitorear los comportamientos de estas distancias durante más tiempo.

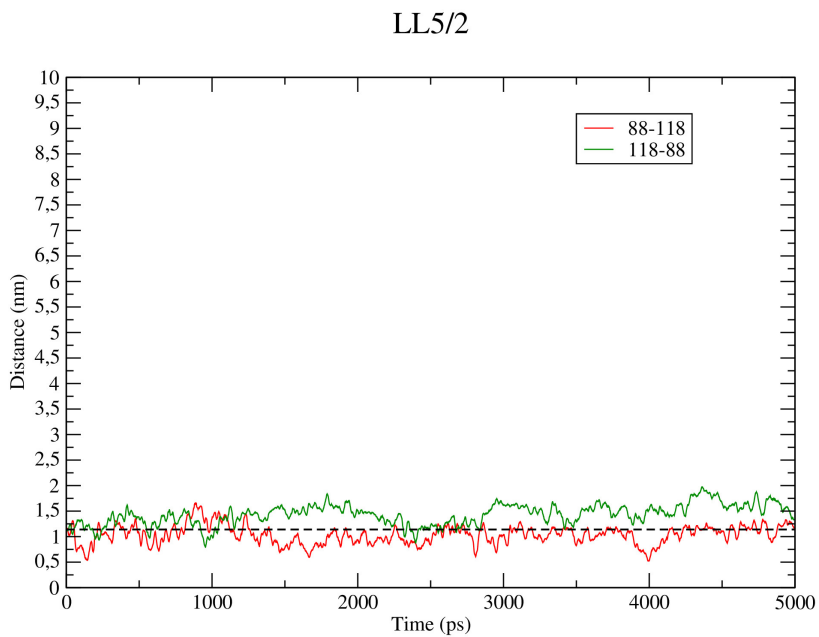


Figura 4.17:
Análisis de distancias en los 5 ns de simulación del modelo LL5/2 entre los N
de las lisinas 88-118 y 118-88.

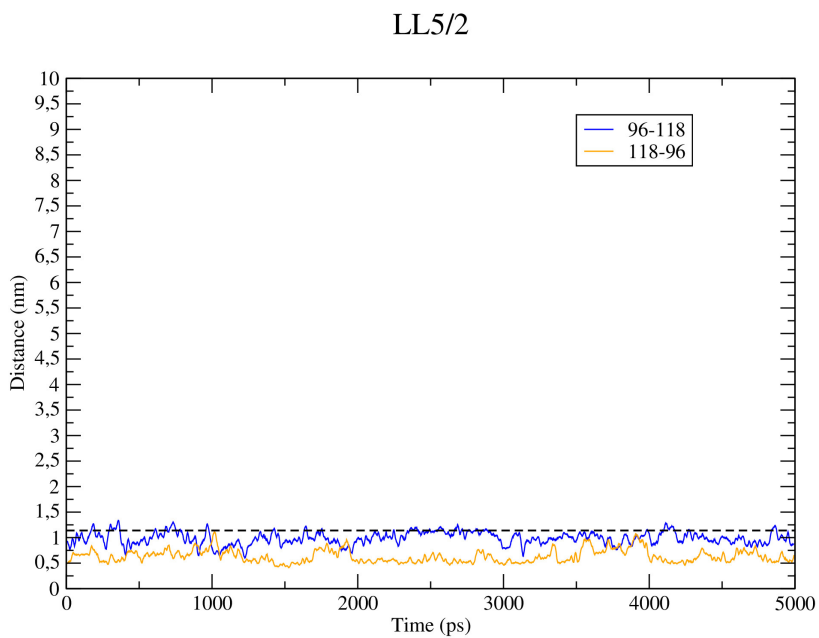


Figura 4.18:
Análisis de distancias en los 5 ns de simulación del modelo LL5/2 entre los N
de las lisinas 96-118 y 118-96.

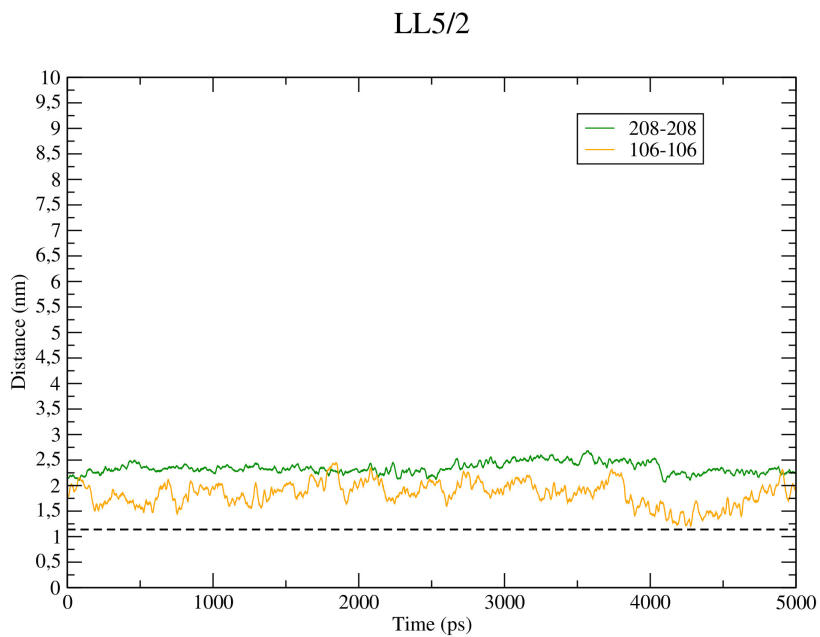


Figura 4.19:
Análisis de distancias en los 5 ns de simulación del modelo LL5/2 entre los N
de las lisinas 208-208 y 106-106.

Capítulo 5

Conclusiones

Ambos modelos de partículas son estables, como se observa a partir de los resultados de los parametros evaluados. En el análisis de superficie accesible a solvente se observan valores constantes para las proteínas y los lipidos, y para el total de la partícula. Estos valores giran en torno a un valor promedio, evidenciando una gran estabilidad. En el análisis de las matrices RMSD, se observan cambios homogéneos de este parámetro, sin que se observe la formación de clusters conformacionales, indicando que la estructura recorre su espacio conformacional de manera gradual, siempre dentro de un mismo estado conformacional. Mediante el análisis de estructura secundaria se observa claramente el predominio de alfa hélices, cuya estructura se mantiene a lo largo de toda la simulación, evidenciando así su estabilidad.

Cuando se comparan los datos experimentales con los obtenidos mediante simulación, se verifica una notable concordancia. En primer lugar, se verifica una predominancia de estructuras de alfa hélices en ambos modelos, como se observa en numerosos experimentos [Mei and Atkinson, 2011, Gursky, 2015]. En comparación con los enlaces por entrecruzamiento propuestos por Silva y otros [Silva et al., 2005], para el modelo de cinturón doble en registro LL5/5 se verifico 3 de los 6 enlaces previstos, y para el modelo en registro LL5/2 se verifico 2 de los 4 enlaces previstos, quedando uno indefinido.

Por lo anteriormente mencionado, la comparación de los resultados entre los dos modelos establece que, el modelo LL5/2 es al menos tan estable como el modelo LL5/5, y que los resultados del modelo LL5/2 son, al menos, comparables con los obtenidos por el modelo LL5/5 frente a los datos experimentales

actualmente disponibles en bibliografía.

Bibliografía

- [Berendsen et al., 1987] Berendsen, H. J. C., Grigera, J. R., and Straatsma, T. P. (1987). The Missing Term in Effective Pair Potentials. *Journal of Physical Chemistry*, 91(24):6269–6271.
- [Berger et al., 1997] Berger, O., Edholm, O., and Jähnig, F. (1997). Molecular dynamics simulations of a fluid bilayer of dipalmitoylphosphatidylcholine at full hydration, constant pressure, and constant temperature. *Biophysical journal*, 72(5):2002–13.
- [Best et al., 2008] Best, R. B., Buchete, N.-V., and Hummer, G. (2008). Are current molecular dynamics force fields too helical? *Biophysical journal*, 95(1):L07–L09.
- [Chiu et al., 2009] Chiu, S. W., Pandit, S. a., Scott, H. L., and Jakobsson, E. (2009). An improved united atom force field for simulation of mixed lipid bilayers. *Journal of Physical Chemistry B*, 113:2748–2763.
- [Eargle et al., 2006] Eargle, J., Wright, D., and Luthey-Schulten, Z. (2006). Multiple Alignment of protein structures and sequences for VMD. *Bioinformatics*, 22(4):504–506.
- [Eisenhaber et al., 1995] Eisenhaber, F., Lijnzaad, P., Argos, P., Sander, C., and Scharf, M. (1995). The double cubic lattice method: Efficient approaches to numerical integration of surface area and volume and to dot surface contouring of molecular assemblies. *Journal of Computational Chemistry*, 16(3):273–284.
- [Feingold and Grunfeld, 2015] Feingold, K. R. and Grunfeld, C. (2015). Introduction to Lipids and Lipoproteins.

- [Frishman and Argos, 1995] Frishman, D. and Argos, P. (1995). Knowledge-based secondary structure assignment. *Proteins: structure, function and genetics*, 23:566–579.
- [Gogonea, 2016] Gogonea, V. (2016). Structural Insights into High Density Lipoprotein: Old Models and New Facts. *Frontiers in Pharmacology*, 6(January).
- [Guha et al., 2008] Guha, M., Gao, X., Jayaraman, S., and Gursky, O. (2008). Correlation of Structural Stability with Functional Remodeling of High-Density Lipoproteins : The Importance of Being Disordered †. pages 11393–11397.
- [Gursky, 2015] Gursky, O. (2015). Structural stability and functional remodeling of high-density lipoproteins. *FEBS Letters*.
- [Hub et al., 2010] Hub, J. S., Winkler, F. K., Merrick, M., and de Groot, B. L. (2010). Potentials of mean force and permeabilities for carbon dioxide, ammonia, and water flux across a Rhesus protein channel and lipid membranes. *Journal of the American Chemical Society*, 132(38):13251–63.
- [Humphrey et al., 1996] Humphrey, W., Dalke, A., and Schulten, K. (1996). {VMD} – {V}isual {M}olecular {D}ynamics. *Journal of Molecular Graphics*, 14:33–38.
- [Jones et al., 2009] Jones, M. K., Catta, A., Patterson, J. C., Gu, F., Chen, J., Li, N., and Segrest, J. P. (2009). Thermal stability of apolipoprotein A-I in high-density lipoproteins by molecular dynamics. *Biophysical Journal*, 96(2):354–371.
- [Kwiterovich, 2000] Kwiterovich, P. O. (2000). The metabolic pathways of high-density lipoprotein, low-density lipoprotein, and triglycerides: a current review. *The American journal of cardiology*, 86(12A):5L–10L.
- [Lagerstedt et al., 2012] Lagerstedt, J. O., Budamagunta, M. S., Liu, G. S., Devalle, N. C., Voss, J. C., and Oda, M. N. (2012). The "beta-clasp" model of apolipoprotein A-I - A lipid-free solution structure determined by electron paramagnetic resonance spectroscopy. *Biochimica et biophysica acta*, 1821(3):448–55.

- [Lindahl et al.,] Lindahl, E., Hess, B., and van der Spoel, D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. *Molecular modeling annual*, 7(8):306–317.
- [Mei and Atkinson, 2011] Mei, X. and Atkinson, D. (2011). Crystal structure of C-terminal truncated apolipoprotein A-I reveals the assembly of High Density Lipoprotein (HDL) by dimerization. *Journal of Biological Chemistry*, 286(44):38570–38582.
- [Mendis et al., 2011] Mendis, S., Puska, P., and Norrving, B. (2011). Global atlas on cardiovascular disease prevention and control. *World Health Organization*, pages 2–14.
- [Miller, 1987] Miller, N. E. (1987). Associations of high-density lipoprotein subclasses and apolipoproteins with ischemic heart disease and coronary atherosclerosis. *American heart journal*, 113(2 Pt 2):589–597.
- [Phillips, 2013] Phillips, M. C. (2013). New insights into the determination of HDL structure by apolipoproteins: Thematic review series: high density lipoprotein structure, function, and metabolism. *Journal of lipid research*, 54(8):2034–48.
- [Sanner et al., 1995] Sanner, M., Olsen, A., and Spehner, J.-C. (1995). Fast and Robust Computation of Molecular Surfaces. In *Proceedings of the 11th ACM Symposium on Computational Geometry*, pages C6–C7, New York. ACM.
- [Scott et al., 1999] Scott, W. R. P., Hünenberger, P. H., Tironi, I. G., Mark, A. E., Billeter, S. R., Fennen, J., Torda, A. E., Huber, T., Krüger, P., and van Gunsteren, W. F. (1999). The GROMOS Biomolecular Simulation Program Package. *The Journal of Physical Chemistry A*, 103(19):3596–3607.
- [Segrest et al., 2000] Segrest, J. P., Harvey, S. C., and Zannis, V. (2000). Detailed molecular model of apolipoprotein A-I on the surface of high-density lipoproteins and its functional implications. *Trends in Cardiovascular Medicine*, 10(6):246–252.
- [Segrest et al., 1999] Segrest, J. P., Jones, M. K., Klon, A. E., Sheldahl, C. J., Hellinger, M., De Loof, H., and Harvey, S. C. (1999). A detailed molecular belt model for apolipoprotein A-I in discoidal high density lipoprotein. *Journal of Biological Chemistry*, 274(45):31755–31758.

- [Sharma et al., 2000] Sharma, R., Zeller, M., Pavlovic, V. I., Huang, T. S., Lo, Z., Chu, S., Zhao, Y., Phillips, J. C., and Schulten, K. (2000). Speech/Gesture Interface to a Visual-Computing Environment. 20:29–37.
- [Silva et al., 2005] Silva, R. A. G. D., Hilliard, G. M., Li, L., Segrest, J. P., and Davidson, W. S. (2005). A mass spectrometric determination of the conformation of dimeric apolipoprotein A-I in discoidal high density lipoproteins. *Biochemistry*, 44(24):8600–7.
- [Sommer et al., 2011a] Sommer, B., Dingersen, T., Gamroth, C., Heissmann, A. J., Lukat, G., Rotzoll, R., Rubert, S., Schäfer, A., and Krüger, J. (2011a). Coarse-grained and all-atom MD simulations with Gromacs based on CELLmicrocosmos 2.2 model membranes.
- [Sommer et al., 2011b] Sommer, B., Dingersen, T., Gamroth, C., Schneider, S. E., Rubert, S., Kruger, J., and Dietz, K.-J. (2011b). CELLmicrocosmos 2.2 MembraneEditor: A Modular Interactive Shape-Based Software Approach To Solve Heterogeneous Membrane Packing Problems. *Journal of chemical information and modeling*, pages 1165–1182.
- [Stone, 1998] Stone, J. (1998). \em An Efficient Library for Parallel Ray Tracing and Animation. Master’s thesis, Computer Science Department, University of Missouri-Rolla.
- [Stone et al., 2001] Stone, J., Gullingsrud, J., Grayson, P., and Schulten, K. (2001). A System for Interactive Molecular Dynamics Simulation. In Hughes, J. F and Séquin, C. H., editors, *2001 ACM Symposium on Interactive 3D Graphics*, pages 191–194, New York. ACM SIGGRAPH.
- [van Gunsteren et al., 1996] van Gunsteren, W., Billeter, S., Eising, A., Hünenberger, P., Krüger, P., Mark, A., Scott, W., and Tironi, I. (1996). Biomolecular Simulation: The {GROMOS96} manual and user guide.
- [Varshney et al., 1994] Varshney, A., Brooks, F. P., and Wright, W. V. (1994). Linearly Scalable Computation of Smooth Molecular Surfaces. *IEEE Computer Graphics and Applications*, 14:19–25.
- [Vergeer et al., 2010] Vergeer, M., Holleboom, A. G., Kastelein, J. J. P., and Kuivenhoven, J. A. (2010). The HDL hypothesis: does high-density lipoprotein protect from atherosclerosis? *Journal of lipid research*, 51(8):2058–2073.

- [Wennberg et al., 2012] Wennberg, C. L., van der Spoel, D., and Hub, J. S. (2012). Large influence of cholesterol on solute partitioning into lipid membranes. *Journal of the American Chemical Society*, 134(11):5351–61.
- [Wilson et al., 1988] Wilson, P. W., Abbott, R. D., and Castelli, W. P. (1988). High density lipoprotein cholesterol and mortality. The Framingham Heart Study. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 8(6):737–741.
- [Wu et al., 2007] Wu, Z., Wagner, M. a., Zheng, L., Parks, J. S., Shy, J. M., Smith, J. D., Gogonea, V., and Hazen, S. L. (2007). The refined structure of nascent HDL reveals a key functional domain for particle maturation and dysfunction. *Nature structural & molecular biology*, 14(9):861–868.